

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 1, 2002, 06:39:49 ; Search time 1821.26 Seconds
(without alignments)
6181.697 Million cell updates/sec

Title: US-09-782-745-14
Perfect score: 538
Sequence: 1 ACGCCAGGAGCTGTGAGGC.....CTGCAAGAAAAA 538

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues
Total number of hits satisfying chosen parameters: 2011660

Minimum DB seq length: 0
Maximum DB seq length: 538

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sv:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htgo_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
------------	-------	-------	--------	----	-------------

RESULT	1	ALIGNMENTS	538 bp	DNA	linear	PAT 29-SEP-1999
AR028488	AR028488	Sequence 14	from patent US 5858689.			
LOCUS	AR028488	Sequence 14	from patent US 5858689.			
DEFINITION	AR028488	Sequence 14	from patent US 5858689.			
ACCESSION	AR028488	Sequence 14	from patent US 5858689.			
VERSION	AR028488.1	GI:5940461				
KEYWORDS						
SOURCE						
ORGANISM						
REFERENCE						
AUTHORS						
TITLE						
JOURNAL						
FEATURES						
BASE COUNT						
ORIGIN						

Query Match 100.0%; Score 538; DB 6; Length 538; Best Local Similarity 100.0%; Pred. No. 6.6e-120; Matches 538; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	ACGCCAGGGAGCTGTGAGGCGAGTGTGTGGTTCCTGCGGTCCGGACTCTTTTTCCTCT	60						
Db	1	ACGCCAGGGAGCTGTGAGGCGAGTGTGTGGTTCCTGCGGTCCGGACTCTTTTTCCTCT	60						
QY	61	ACTGAGATTTCATCTGTGTGAATATGAGTTGGCGAGGAAGATCGACCTATCGGCCTAGAC	120						
Db	61	ACTGAGATTTCATCTGTGTGAATATGAGTTGGCGAGGAAGATCGACCTATCGGCCTAGAC	120						
QY	121	CAGACGCTACGTAGAGCCTCCTGAAATGATTGGGCTATGCGGCCGAGCAGTTCAAGT	180						
Db	121	CAGACGCTACGTAGAGCCTCCTGAAATGATTGGGCTATGCGGCCGAGCAGTTCAAGT	180						
QY	181	ATGAAGTGGAAACCAAGACCTGGAAGAGGGGAACCCAGCAACTCAACGTCAGGATCCTG	240						
Db	181	ATGAAGTGGAAACCAAGACCTGGAAGAGGGGAACCCAGCAACTCAACGTCAGGATCCTG	240						
QY	241	CAGTGTCTCAGGAGGGAGGAGTGAAGGAGCATCTCGAGTCAAGGCCGAAAGCCTGAAG	300						
Db	241	CAGTGTCTCAGGAGGGAGGAGTGAAGGAGCATCTCGAGTCAAGGCCGAAAGCCTGAAG	300						
QY	301	CTCATAGCCAGAACAGGTCACCCACAGACTGGGTGGTGTGAGTGTGAAGATGGTCTGATG	360						
Db	301	CTCATAGCCAGAACAGGTCACCCACAGACTGGGTGGTGTGAGTGTGAAGATGGTCTGATG	360						
QY	361	GGCAGGAGATGGAGCCGCAATCCAGAGGAGGTGAAAACGCCCTGAAGAAGGTGAAAAAGC	420						
Db	361	GGCAGGAGATGGAGCCGCAATCCAGAGGAGGTGAAAACGCCCTGAAGAAGGTGAAAAAGC	420						
QY	421	AATCACAGTGTAAAGAAGACACGTTGAAATGATGCGAGCTCTCTATGTTGGAATTT	480						
Db	421	AATCACAGTGTAAAGAAGACACGTTGAAATGATGCGAGCTCTCTATGTTGGAATTT	480						
QY	481	TGTTTCATTAATAATCTCCCAATAAGCTTTACAGCCTTCTGCAAGAAAAA	538						
Db	481	TGTTTCATTAATAATCTCCCAATAAGCTTTACAGCCTTCTGCAAGAAAAA	538						
RESULT 2									
LOCUS I55851 538 bp DNA linear PAT 07-OCT-1997									
DEFINITION Sequence 14 from patent US 5648226.									
ACCESSION I55851									
VERSION I55851.1 GI:2476645									
KEYWORDS Unknown.									
SOURCE Unknown.									
ORGANISM Unclassified.									
REFERENCE 1 (bases 1 to 538)									
AUTHORS Van den Eynde,B., DeBacker,O. and Boon-Palleur,T.									
TITLE Isolated peptides derived from tumor rejection antigens, and their use									
JOURNAL Patent: US 5648226-A 14 15-JUL-1997;									
FEATURES Location/Qualifiers									
source 1..538									
BASE COUNT 160 a 116 c 155 g 107 t									
ORIGIN									
Query Match 100.0%; Score 538; DB 6; Length 538; Best Local Similarity 100.0%; Pred. No. 6.6e-120; Matches 538; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	ACGCCAGGAGCTGTGAGGCGAGTGTGTGGTTCCTGCGGTCCGGACTCTTTTTCCTCT	60						
Db	1	ACGCCAGGAGCTGTGAGGCGAGTGTGTGGTTCCTGCGGTCCGGACTCTTTTTCCTCT	60						
QY	61	ACTGAGATTTCATCTGTGTGAATATGAGTTGGCGAGGAAGATCGACCTATCGGCCTAGAC	120						
Db	61	ACTGAGATTTCATCTGTGTGAATATGAGTTGGCGAGGAAGATCGACCTATCGGCCTAGAC	120						
QY	121	CAGACGCTACGTAGAGCCTCCTGAAATGATTGGGCTATGCGGCCGAGCAGTTCAAGT	180						
Db	121	CAGACGCTACGTAGAGCCTCCTGAAATGATTGGGCTATGCGGCCGAGCAGTTCAAGT	180						
QY	181	ATGAAGTGGAAACCAAGACCTGGAAGAGGGGAACCCAGCAACTCAACGTCAGGATCCTG	240						
Db	181	ATGAAGTGGAAACCAAGACCTGGAAGAGGGGAACCCAGCAACTCAACGTCAGGATCCTG	240						
QY	241	CAGTGTCTCAGGAGGGAGGAGTGAAGGAGCATCTCGAGTCAAGGCCGAAAGCCTGAAG	300						
Db	241	CAGTGTCTCAGGAGGGAGGAGTGAAGGAGCATCTCGAGTCAAGGCCGAAAGCCTGAAG	300						
QY	301	CTCATAGCCAGAACAGGTCACCCACAGACTGGGTGGTGTGAGTGTGAAGATGGTCTGATG	360						
Db	301	CTCATAGCCAGAACAGGTCACCCACAGACTGGGTGGTGTGAGTGTGAAGATGGTCTGATG	360						
QY	361	GGCAGGAGATGGAGCCGCAATCCAGAGGAGGTGAAAACGCCCTGAAGAAGGTGAAAAAGC	420						
Db	361	GGCAGGAGATGGAGCCGCAATCCAGAGGAGGTGAAAACGCCCTGAAGAAGGTGAAAAAGC	420						
QY	421	AATCACAGTGTAAAGAAGACACGTTGAAATGATGCGAGCTCTCTATGTTGGAATTT	480						
Db	421	AATCACAGTGTAAAGAAGACACGTTGAAATGATGCGAGCTCTCTATGTTGGAATTT	480						
QY	481	TGTTTCATTAATAATCTCCCAATAAGCTTTACAGCCTTCTGCAAGAAAAA	538						
Db	481	TGTTTCATTAATAATCTCCCAATAAGCTTTACAGCCTTCTGCAAGAAAAA	538						
RESULT 3									
LOCUS HSU19143 530 bp mRNA linear PRI 04-DEC-1995									
DEFINITION Human GAGE-2 protein mRNA, complete cds.									
ACCESSION U19143									
VERSION U19143.1 GI:914900									
KEYWORDS human.									
SOURCE Homo sapiens									
ORGANISM Homo sapiens									
REFERENCE 1 (bases 1 to 530)									
AUTHORS Van den Eynde,B., Peeters,O., De Backer,O., Gaugler,B., Lucas,S. and Boon,T.									
TITLE A new family of genes coding for an antigen recognized by autologous cytolytic T lymphocytes on a human melanoma									
JOURNAL J. Exp. Med. 182 (3), 689-698 (1995)									
MEDLINE 95378788									
REFERENCE 2 (bases 1 to 530)									
AUTHORS Van den Eynde,B.J.									
TITLE Direct Submission									
JOURNAL Submitted (28-DEC-1994) Benoit J Van Den Eynde, Ludwig Institute For Cancer Research, 74 Avenue Hippocrate, BRUSSELS, 1200, BELGIUM									
FEATURES Location/Qualifiers									
source 1..530									
/organism="Homo sapiens"									
/db_xref="taxon:9606"									
/sex="female"									
/cell_line="M2-MEL.43"									
/tissue_type="melanoma"									
/dev_stage="adult"									
84..434									
/codon_start=1									
/product="GAGE-2 protein"									
/protein_id="AA82745.1"									
/db_xref="GI:914901"									
/translation="MSMRGRSTYRPRRVRVPEPMIGMRPEQFSDEVEPATPEEGE PATQRDPAAQEGDEGASAGQPKPEAHSEQCHPQTGCECEDGPDQEMDPPNPE EVKTPEEGKQSOC"									
BASE COUNT 152 a 116 c 155 g 107 t									
ORIGIN									



Query Match 98.5%; Score 530; DB 9; Length 530;
Best Local Similarity 100.0%; Pred. No. 5.7e-118; Indels 0; Gaps 0;
Matches 530; Conservative 0; Mismatches 0;

QY 1 ACGCCAGGGAGCTGTGAGCAGCTGTGTGTTCTGCGCTCCGACTCTTTTCTCT 60
Db 1 ACGCCAGGGAGCTGTGAGCAGCTGTGTGTTCTGCGCTCCGACTCTTTTCTCT 60

QY 61 ACTGAGATTCTATCTGTGTAATATAGTTGGGAGGAAGATCGACCTATCGGCTAGAC 120
Db 61 ACTGAGATTCTATCTGTGTAATATAGTTGGGAGGAAGATCGACCTATCGGCTAGAC 120

QY 121 CAAGACGCTACGTAGAGCTCTGAAATGATTGGGCTTATGCGGCCGAGCAGTTCACTG 180
Db 121 CAAGACGCTACGTAGAGCTCTGAAATGATTGGGCTTATGCGGCCGAGCAGTTCACTG 180

QY 181 ATGAAGTGGAAACAGCAACACCTGAAGAGGGGAACAGCAACTCAACGTCAGATCCTG 240
Db 181 ATGAAGTGGAAACAGCAACACCTGAAGAGGGGAACAGCAACTCAACGTCAGATCCTG 240

QY 241 CAGCTGCTCAGAGGAGGAGATGAGGAGCATCTGCAGGTCAAGGCCGAGGCTGAAG 300
Db 241 CAGCTGCTCAGAGGAGGAGATGAGGAGCATCTGCAGGTCAAGGCCGAGGCTGAAG 300

QY 301 CTCATAGCCAGGAACAGGGTCAACCCACAGACTGGGTGTGAGTGTGAAGTGGTCTGATG 360
Db 301 CTCATAGCCAGGAACAGGGTCAACCCACAGACTGGGTGTGAGTGTGAAGTGGTCTGATG 360

QY 361 GCGAGGAGATGACCCGCCCAATCCAGAGGAGGTGAAGCCCTGAAGAGGTGAAGAGC 420
Db 361 GCGAGGAGATGACCCGCCCAATCCAGAGGAGGTGAAGCCCTGAAGAGGTGAAGAGC 420

QY 421 AATCACAGTGTAAAGAGACACGCTTGAATATGATGAGGCTGCTCTATGTTGAAATT 480
Db 421 AATCACAGTGTAAAGAGACACGCTTGAATATGATGAGGCTGCTCTATGTTGAAATT 480

QY 481 TGTTCATTAAATTTCTCCCAATAAAGCTTTACAGCCTTCGCAAGAAAA 530
Db 481 TGTTCATTAAATTTCTCCCAATAAAGCTTTACAGCCTTCGCAAGAAAA 530

RESULT 4
AF055473 AF055473 528 bp mRNA linear PRI 01-MAY-2000
LOCUS Homo sapiens GAGE-8 mRNA, complete cds.
DEFINITION AF055473
ACCESSION AF055473.1 GI:3511022
VERSION
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 528)
AUTHORS De Backer, O., Arden, K.C., Boretti, M., Vantomme, V., De Smet, C., Czekay, S., Vliers, C.S., De Plaen, E., Brasseur, F., Chomez, P., Van den Eynde, B., Boon, T. and van der Bruggen, P.
TITLE Characterization of the GAGE genes that are expressed in various human cancers and in normal testis
JOURNAL Cancer Res. 59 (13), 3157-3165 (1999)
MEDLINE 99323388
PUBMED 10397259
REFERENCE 2 (bases 1 to 528)
AUTHORS De Backer, O.R.Y.
TITLE Direct Submission
JOURNAL Submitted (24-MAR-1998) Ludwig Institute for Cancer Research, Brussels Branch, 74 av. Hippocrate, Brussels B-1200, Belgium
FEATURES
source
1..528
/organism="Homo sapiens"
/db_xref="taxon:9606"
73..423
/codon_start=1

/product="GAGE-8"
/protein_id="AAC33676.1"
/db_xref="GI:3511023"
/translation="MSWRGRSTYRPRPRRYVEPMIGPMRPEOFSDVEPATPEEGE
PATQRPAAQEGEDGASAGQPKPEADSBQGHPTQCECEDGPDQEMDPNPE
EVKTPDEGEKQSQ"

BASE COUNT 158 a 112 c 151 g 107 t
ORIGIN

Query Match 97.7%; Score 525.4; DB 9; Length 528;
Best Local Similarity 99.8%; Pred. No. 7.3e-117; Indels 0; Gaps 0;
Matches 526; Conservative 0; Mismatches 1;

QY 12 CTGTGAGCAGCTGCTGTGTGTTCTGCGCTCCGACTCTTTTCTCTACTGAGATTCA 71
Db 1 CTGTGAGCAGCTGCTGTGTGTTCTGCGCTCCGACTCTTTTCTCTACTGAGATTCA 60

QY 72 TCTGTGTAAATATAGTTGGGAGGAAGATCGACCTATCGGCCCTAGACCAAGCGCTAC 131
Db 61 TCTGTGTAAATATAGTTGGGAGGAAGATCGACCTATCGGCCCTAGACCAAGCGCTAC 120

QY 132 GTAGAGCCTCTGAAATGATTGGCCCTATCGGCCCGAGCAGTTCAGTGAAGTGGAA 191
Db 121 GTAGAGCCTCTGAAATGATTGGCCCTATCGGCCCGAGCAGTTCAGTGAAGTGGAA 180

QY 192 CCAGCAACACCTGAAAGAGGGGAACCAAGCAACTCAAGTCAAGTCTGCAGCTGCTCAG 251
Db 181 CCAGCAACACCTGAAAGAGGGGAACCAAGCAACTCAAGTCAAGTCTGCAGCTGCTCAG 240

QY 252 GAGGAGAGGATGAGGAGCATCTGCAGGTCAAGGCCGGAAGCCTGAAGCTCATAGCCAG 311
Db 241 GAGGAGAGGATGAGGAGCATCTGCAGGTCAAGGCCGGAAGCCTGAAGCTCATAGCCAG 300

QY 312 GAACAGGCTCAACCCACAGACTGGGTGTGAGTGTGAAGATGCTCTGTGGCAGGAGATG 371
Db 301 GAACAGGCTCAACCCACAGACTGGGTGTGAGTGTGAAGATGCTCTGTGGCAGGAGATG 360

QY 372 GACCCGCCAATCCAGAGGAGTGAAGCCGCTGAAGAGGTGAAGCAATTCACAGTGT 431
Db 361 GACCCGCCAATCCAGAGGAGTGAAGCCGCTGAAGAGGTGAAGCAATTCACAGTGT 420

QY 432 TAAAGAGACACAGCTTGAATGATGACGCTGCTCTATGTTGGAAATTTGTTCAATTA 491
Db 421 TAAAGAGACACAGCTTGAATGATGACGCTGCTCTATGTTGGAAATTTGTTCAATTA 480

QY 492 ATCTCTCCCAATAAAGCTTTACAGCCTTCGCAAGAAAAA 538
Db 481 ATCTCTCCCAATAAAGCTTTACAGCCTTCGCAAGAAAAA 527

RESULT 5
AR028491 AR028491 532 bp DNA linear PAT 29-SEP-1999
LOCUS Sequence 17 from patent US 5858689.
DEFINITION AR028491
ACCESSION AR028491
VERSION AR028491.1 GI:5940464
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 532)
AUTHORS van der Bruggen, P., van den Eynde, B., DeBacker, O. and Boon-Falleur, T.
TITLE Isolated peptides derived from the gage tumpr rejection antigen precursor and uses thereof
JOURNAL Patent: US 5858689-A 17 12-JAN-1999;
FEATURES
source
1..532
/organism="unknown"
156 a 111 c 154 g 111 t
ORIGIN

Query Match 93.8%; Score 504.8; DB 6; Length 532;
Best Local Similarity 98.1%; Pred. No. 7.1e-112;
Matches 522; Conservative 0; Mismatches 7; Indels 3; Gaps 1;

QY 10 AGCTGTGAGGCGAGTGTGTGGTTCCTCGCGTCCGGACTCTTTTCTCTACTGAGATT 69
|||||
DB 1 AGCTGTGAGGCGAGTGTGTGGTTCCTCGCGTCCGGACTCTTTTCTCTACTGAGATT 60
|||||

QY 70 CATCTGTGTGAATATGAGTTGGCGAGGAAGATCGACCC---TATCGGCGCTAGACCAAGAC 126
|||||
DB 61 CATCTGTGTGAATATGAGTTGGCGAGGAAGATCGACCTATTATTGGCTTAGACCAAGGC 120
|||||

QY 127 GCTAGTGTAGAGCCTCCTGAAATGATTGGGCTTATCGCGCCGAGCAGTTCAGTGATGAAG 186
|||||
DB 121 GCTATGTACAGCCTCCTGAAGTGATTGGGCTTATCGCGCCGAGCAGTTCAGTGATGAAG 180
|||||

QY 187 TGGACCAAGCAACCTGTAAGAGGGGAAACCAAGCTCAACGCTCAGGATCCTGCAGCTG 246
|||||
DB 181 TGGACCAAGCAACCTGTAAGAGGGGAAACCAAGCTCAACGCTCAGGATCCTGCAGCTG 240
|||||

QY 247 CTCAGGGGAGAGGATGAGGAGCATCTCGAGTCAAGGCGCGAGCAGTTCAGTGATGAAG 306
|||||
DB 241 CTCAGGGGAGAGGATGAGGAGCATCTCGAGTCAAGGCGCGAGCAGTTCAGTGATGAAG 300
|||||

QY 307 GCCAGGAACAGGTCACCCACAGACTGGGTGTGAGTGTGAAGATGGTCTCGATGGGCAGG 366
|||||
DB 301 GCCAGGAACAGGTCACCCACAGACTGGGTGTGAGTGTGAAGATGGTCTCGATGGGCAGG 360
|||||

QY 367 AGATGGACCCGCCAAATCCAGAGGAGTGAACGCTTCAAGGCGCGAGCAGTTCAGTGATGAAG 426
|||||
DB 361 AGATGGACCCGCCAAATCCAGAGGAGTGAACGCTTCAAGGCGCGAGCAGTTCAGTGATGAAG 420
|||||

QY 427 AGTGTAAAAGAGACAGCTTGAAATGATGCAGGCTGCTCTATGTTGGAAATTTGTTCA 486
|||||
DB 421 AGTGTAAAAGAGACAGCTTGAAATGATGCAGGCTGCTCTATGTTGGAAATTTGTTCA 480
|||||

QY 487 TTAATAATCTCCCAATAAGCTTTTACAGCCTTCTGCAAGAAAAA 538
|||||
DB 481 TTAATAATCTCCCAATAAGCTTTTACAGCCTTCTGCAAGAAAAA 532
|||||

RESULT 6
LOCUS I55854
DEFINITION Sequence 17 from patent US 5648226.
ACCESSION I55854
VERSION I55854.1 GI:2476648
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 532)
AUTHORS Van den Eynde,B., DeBacker,O. and Boon-Falleur,T.
TITLE Isolated peptides derived from tumor rejection antigens, and their use
JOURNAL Patent: US 5648226-A 17 15-JUL-1997;
FEATURES Location/Qualifiers
source
1. .532
BASE COUNT 156 a 111 c 154 g 111 t
ORIGIN

Query Match 93.8%; Score 504.8; DB 6; Length 532;
Best Local Similarity 98.1%; Pred. No. 7.1e-112;
Matches 522; Conservative 0; Mismatches 7; Indels 3; Gaps 1;

QY 10 AGCTGTGAGGCGAGTGTGTGGTTCCTCGCGTCCGGACTCTTTTCTCTACTGAGATT 69
|||||
DB 1 AGCTGTGAGGCGAGTGTGTGGTTCCTCGCGTCCGGACTCTTTTCTCTACTGAGATT 60
|||||

QY 70 CATCTGTGTGAATATGAGTTGGCGAGGAAGATCGACCC---TATCGGCGCTAGACCAAGAC 126
|||||
DB 61 CATCTGTGTGAATATGAGTTGGCGAGGAAGATCGACCTATTATTGGCTTAGACCAAGGC 120
|||||

QY 127 GCTAGTGTAGAGCCTCCTGAAATGATTGGGCTTATCGCGCCGAGCAGTTCAGTGATGAAG 186
|||||
DB 121 GCTATGTACAGCCTCCTGAAGTGATTGGGCTTATCGCGCCGAGCAGTTCAGTGATGAAG 180
|||||

QY 187 TGGACCAAGCAACCTGTAAGAGGGGAAACCAAGCTCAACGCTCAGGATCCTGCAGCTG 246
|||||
DB 181 TGGACCAAGCAACCTGTAAGAGGGGAAACCAAGCTCAACGCTCAGGATCCTGCAGCTG 240
|||||

QY 247 CTCAGGGGAGAGGATGAGGAGCATCTCGAGTCAAGGCGCGAGCAGTTCAGTGATGAAG 306
|||||
DB 241 CTCAGGGGAGAGGATGAGGAGCATCTCGAGTCAAGGCGCGAGCAGTTCAGTGATGAAG 300
|||||

QY 307 GCCAGGAACAGGTCACCCACAGACTGGGTGTGAGTGTGAAGATGGTCTCGATGGGCAGG 366
|||||
DB 301 GCCAGGAACAGGTCACCCACAGACTGGGTGTGAGTGTGAAGATGGTCTCGATGGGCAGG 360
|||||

QY 367 AGATGGACCCGCCAAATCCAGAGGAGTGAACGCTTCAAGGCGCGAGCAGTTCAGTGATGAAG 426
|||||
DB 361 AGATGGACCCGCCAAATCCAGAGGAGTGAACGCTTCAAGGCGCGAGCAGTTCAGTGATGAAG 420
|||||

QY 427 AGTGTAAAAGAGACAGCTTGAAATGATGCAGGCTGCTCTATGTTGGAAATTTGTTCA 486
|||||
DB 421 AGTGTAAAAGAGACAGCTTGAAATGATGCAGGCTGCTCTATGTTGGAAATTTGTTCA 480
|||||

QY 487 TTAATAATCTCCCAATAAGCTTTTACAGCCTTCTGCAAGAAAAA 538
|||||
DB 481 TTAATAATCTCCCAATAAGCTTTTACAGCCTTCTGCAAGAAAAA 532
|||||

Db 61 CATCTGTGTGAATATGAGTTGGCGAGGAAGATCGACCTATTATTGGCTTAGACCAAGGC 120
|||||

QY 127 GCTAGTGTAGAGCCTCCTGAAATGATTGGGCTTATCGCGCCGAGCAGTTCAGTGATGAAG 186
|||||

Db 121 GCTATGTACAGCCTCCTGAAGTGATTGGGCTTATCGCGCCGAGCAGTTCAGTGATGAAG 180
|||||

QY 187 TGGACCAAGCAACCTGTAAGAGGGGAAACCAAGCTCAACGCTCAGGATCCTGCAGCTG 246
|||||

Db 181 TGGACCAAGCAACCTGTAAGAGGGGAAACCAAGCTCAACGCTCAGGATCCTGCAGCTG 240
|||||

QY 247 CTCAGGGGAGAGGATGAGGAGCATCTCGAGTCAAGGCGCGAGCAGTTCAGTGATGAAG 306
|||||

Db 241 CTCAGGGGAGAGGATGAGGAGCATCTCGAGTCAAGGCGCGAGCAGTTCAGTGATGAAG 300
|||||

QY 307 GCCAGGAACAGGTCACCCACAGACTGGGTGTGAGTGTGAAGATGGTCTCGATGGGCAGG 366
|||||

Db 301 GCCAGGAACAGGTCACCCACAGACTGGGTGTGAGTGTGAAGATGGTCTCGATGGGCAGG 360
|||||

QY 367 AGATGGACCCGCCAAATCCAGAGGAGTGAACGCTTCAAGGCGCGAGCAGTTCAGTGATGAAG 426
|||||

Db 361 AGATGGACCCGCCAAATCCAGAGGAGTGAACGCTTCAAGGCGCGAGCAGTTCAGTGATGAAG 420
|||||

QY 427 AGTGTAAAAGAGACAGCTTGAAATGATGCAGGCTGCTCTATGTTGGAAATTTGTTCA 486
|||||

Db 421 AGTGTAAAAGAGACAGCTTGAAATGATGCAGGCTGCTCTATGTTGGAAATTTGTTCA 480
|||||

QY 487 TTAATAATCTCCCAATAAGCTTTTACAGCCTTCTGCAAGAAAAA 538
|||||

Db 481 TTAATAATCTCCCAATAAGCTTTTACAGCCTTCTGCAAGAAAAA 532
|||||

RESULT 7
AX334151
LOCUS AX334151
DEFINITION Sequence 4660 from Patent WO0194629.
ACCESSION AX334151
VERSION AX334151.1 GI:18124870
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (sites)
AUTHORS Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endress,G.,
Horrihan,S., Soppet,D.R. and Weaver,Z.
TITLE Cancer gene determination and therapeutic screening using signature
JOURNAL Patent: WO 0194629-A 4660 13-DEC-2001;
FEATURES Location/Qualifiers
source
1. .528
BASE COUNT 147 a 114 c 156 g 111 t
ORIGIN

Query Match 93.2%; Score 501.4; DB 6; Length 528;
Best Local Similarity 98.3%; Pred. No. 4.7e-111;
Matches 518; Conservative 0; Mismatches 6; Indels 3; Gaps 1;

QY 2 CGCCAGGAGAGCTGTGAGGAGTGTGTGGTTCCTGCGTCCGGACTCTTTTTCCTCTA 61
|||||

Db 1 CGCCAGGAGAGCTGTGAGGAGTGTGTGGTTCCTGCGTCCGGACTCTTTTTCCTCTA 60
|||||

QY 62 CTGAGATTTCATCTGTGTGAATATGAGTTGGCGAGGAAGATCGACCCATATTGGCCTAG 118
|||||

Db 61 CTGAGATTTCATCTGTGTGAATATGAGTTGGCGAGGAAGATCGACCTATTATTGGCCTAG 120
|||||

QY 119 ACCAAGCAGCTAGCTAGAGCCTCTCGAAATGATTGGGCTTATGGCGCCGAGCAGTTCAG 178
|||||

Db 121 ACCAAGCAGCTAGCTAGAGCCTCTCGAAATGATTGGGCTTATGGCGCCGAGCAGTTCAG 180
|||||


```
Qy 179 TGATGAAGTGAACCAACACACCTGAAGAAGGGAACACCACTCAACGTCAGGATCC 238
|||||
Db 181 TGATGAAGTGAACCAACACACCTGAAGAAGGGAACCACTCAACGTCAGGATCC 240
|||||
Qy 239 TGCAGCTGCTCAGAGGAGAGGATGAGGGAGCATCTCAGGTCGAAGGCCGGAAGCCCTGA 298
|||||
Db 241 TGCAGCTGCTCAGAGGAGAGGATGAGGGAGCATCTCAGGTCGAAGGCCGGAAGCCCTGA 300
|||||
Qy 299 AGCTCATAGCCAGAACAGGTCACCCACAGACTGGGTGTGAGTGTGAAGATGGTCCCTGA 358
|||||
Db 301 AGCTCATAGCCAGAACAGGTCACCCACAGACTGGGTGTGAGTGTGAAGATGGTCCCTGA 360
|||||
Qy 359 TGGCAGAGAGATGAGCCGCCAATCCAGAGAGAGTGAAGCGCTGAAGAGGTGAAGA 418
|||||
Db 361 TGGCAGAGAGATGAGCCGCCAATCCAGAGAGAGTGAAGCGCTGAAGAGGTGAAGA 420
|||||
Qy 419 GCAATCAGAGTGTAAAGAGAGACAGCTTGAATGATGAGGCTGCTCTATGTTGGA 478
|||||
Db 421 GCAATCAGAGTGTAAAGAGAGACAGCTTGAATGATGAGGCTGCTCTATGTTGGA 480
|||||
Qy 479 TTTGTTCAATAAATCTCCCAATAAAGCTTTTACAGCCTTCTGCAAA 525
|||||
Db 481 TTTGTTCAATAAATCTCCCAATAAAGCTTTTACAGCCTTCTGCAAA 527
|||||

RESULT 8
HSU19145 528 bp mRNA linear PRI 04-DEC-1995
LOCUS
DEFINITION Human GAGE-4 protein mRNA, complete cds.
ACCESSION U19145
VERSION U19145.1 GI:914904
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 528)
AUTHORS Van den Eynde,B., Peeters,O., De Backer,O., Gaugler,B., Lucas,S.
and Boon,T.
TITLE A new family of genes coding for an antigen recognized by
autologous cytolytic T lymphocytes on a human melanoma
J. Exp. Med. 182 (3), 689-698 (1995)
MEDLINE 95378788
REFERENCE 2 (bases 1 to 528)
AUTHORS Van den Eynde,B.J.
TITLE Direct Submission
JOURNAL Submitted (28-DEC-1994) BENOIT J VAN DEN EYNDE, Ludwig Institute
for Cancer Research, 74 Avenue Hippocrate, Brussels B-1200, Belgium
FEATURES
Source
1..528
/organism="Homo sapiens"
/db_xref="taxon:9606"
/sex="female"
/cell_line="M22-MEL.43"
/tissue_type="melanoma"
/dev_stage="adult"
83..436
/codon_start=1
/product="GAGE-4 protein"
/protein_id="AAA82747.1"
/db_xref="GI:914905"
CDS
EPATQODPAAQGEDEGASGQPKPADSQEQHPQTGCECEDGPDGQEMDPNP
EVKTPERQSQSC"
BASE COUNT 147 a 114 c 156 g 111 t
ORIGIN

Query Match 93.2%; Score 501.4; DB 9; Length 528;
Best Local Similarity 98.3%; Pred. No. 4.7e-111;
Matches 518; Conservative 0; Mismatches 6; Indels 3; Gaps 1;
```

```
Qy 2 CGCCAGGAGCTGTGAGGCAAGTCTGTGTGTTCTCTCCGCTCCGGAAGTCTTTTCTCTCTA 61
|||||
Db 1 CGCCAGGAGCTGTGAGGCAAGTCTGTGTGTTCTCTCCGCTCCGGAAGTCTTTTCTCTCTA 60
|||||
Qy 62 CTGAGATTCTATCTGTGTGAATATGAGTTGGCGAGGAAGATCGACC--TATCGGCCTAG 118
|||||
Db 61 CTGAGATTCTATCTGTGTGAATATGAGTTGGCGAGGAAGATCGACCATTATTTGGCCCTAG 120
|||||
Qy 119 ACCAAGACGCTACGTAGAGCCTCTGTAATGATTTGGCGCTATGCGGCCGAGCAGTTCAG 178
|||||
Db 121 ACCAAGCGCTATGTACAGCCTCTGTAATGATTTGGCGCTATGCGGCCGAGCAGTTCAG 180
|||||
Qy 179 TGATGAAGTGAACCAACACCTGAAGAGGGAACCACTCAACGTCAGGATCC 238
|||||
Db 181 TGATGAAGTGAACCAACACCTGAAGAGGGAACCACTCAACGTCAGGATCC 240
|||||
Qy 239 TGCAGCTGCTCAGAGGAGGATGAGGGAGCATCTGCGGCTCAAGGCCGGAAGCCCTGA 298
|||||
Db 241 TGCAGCTGCTCAGAGGAGGATGAGGGAGCATCTGCGGCTCAAGGCCGGAAGCCCTGA 300
|||||
Qy 299 AGCTCATAGCCAGAACAGGTCACCCACAGACTGGGTGTGAGTGTGAAGATGGTCCCTGA 358
|||||
Db 301 AGCTCATAGCCAGAACAGGTCACCCACAGACTGGGTGTGAGTGTGAAGATGGTCCCTGA 360
|||||
Qy 359 TGGCAGAGAGATGAGCCGCCAATCCAGAGAGAGTGAAGCGCTGAAGAGGTGAAGA 418
|||||
Db 361 TGGCAGAGAGATGAGCCGCCAATCCAGAGAGAGTGAAGCGCTGAAGAGGTGAAGA 420
|||||
Qy 419 GCAATCAGAGTGTAAAGAGAGACAGCTTGAATGATGAGGCTGCTCTATGTTGGA 478
|||||
Db 421 GCAATCAGAGTGTAAAGAGAGACAGCTTGAATGATGAGGCTGCTCTATGTTGGA 480
|||||
Qy 479 TTTGTTCAATAAATCTCCCAATAAAGCTTTTACAGCCTTCTGCAAA 525
|||||
Db 481 TTTGTTCAATAAATCTCCCAATAAAGCTTTTACAGCCTTCTGCAAA 527
|||||

RESULT 9
HSU19147 527 bp mRNA linear PRI 04-DEC-1995
LOCUS
DEFINITION Human GAGE-6 protein mRNA, complete cds.
ACCESSION U19147
VERSION U19147.1 GI:914908
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 527)
AUTHORS Van den Eynde,B., Peeters,O., De Backer,O., Gaugler,B., Lucas,S.
and Boon,T.
TITLE A new family of genes coding for an antigen recognized by
autologous cytolytic T lymphocytes on a human melanoma
J. Exp. Med. 182 (3), 689-698 (1995)
MEDLINE 95378788
REFERENCE 2 (bases 1 to 527)
AUTHORS Van den Eynde,B.J.
TITLE Direct Submission
JOURNAL Submitted (28-DEC-1994) BENOIT J VAN DEN EYNDE, Ludwig Institute
for Cancer Research, 74 Avenue Hippocrate, Brussels B-1200, Belgium
FEATURES
Source
1..527
/organism="Homo sapiens"
/db_xref="taxon:9606"
/sex="female"
/cell_line="M22-MEL.43"
/tissue_type="melanoma"
/dev_stage="adult"
82..435
/codon_start=1
/product="GAGE-6 protein"
/protein_id="AAA82749.1"
/db_xref="GI:914909"
CDS
```

/translation="MSWRGRSTYYWPRRYVQPEVIGMPRPFQSFDEVEPATPEEG
EPATQRDPAAAGGEGDEGASAGOGPKPEADSBQGHQHPOTGCECEDGPDGQEVDPNP
EEVKTPEGEKQSQ"

BASE COUNT 146 a 113 c 157 g 111 t

ORIGIN

Query Match 92.7%; Score 498.8; DB 9; Length 527;
Best Local Similarity 98.1%; Pred. No. 2e-110;
Matches 516; Conservative 0; Mismatches 7; Indels 3; Gaps 1;

Qy 3 GCCAGGAGCTGTGAGGAGCTGCTGTGTCTCTCGCGTCCGGACTCTTTTCTCTAC 62
Db 1 GCCAGGAGCTGTGAGGAGCTGCTGTGTCTCTCGCGTCCGGACTCTTTTCTCTAC 60
Qy 63 TGAGATCATCTGTGTAATATGAGTGGGAGGAGATCGACC--TATCGGCTAGA 119
Db 61 TGAGATCATCTGTGTAATATGAGTGGGAGGAGATCGACCTATTATGGCTAGA 120
Qy 120 CCAAGAGCTACGTAGAGCTCTCTGAAATGATGGGCTATCGGCGCGGAGCTTCAGT 179
Db 121 CCAAGGCGCTATGTACAGCTCTCTGAGTGTGGGCTATCGGCGCGGAGCTTCAGT 180
Qy 180 GATGAAGTGGAAACAGCAACCTGGAAGAGGGAACAGCAACTCAACGTCAGGATCCT 239
Db 181 GATGAAGTGGAAACAGCAACCTGGAAGAGGGAACAGCAACTCAACGTCAGGATCCT 240
Qy 240 GCAGCTGCTCAGGAGGAGAGGATGAGGAGCATCTGAGGTCAAGGCGCGGAGCTGAA 299
Db 241 GCAGCTGCTCAGGAGGAGAGGATGAGGAGCATCTGAGGTCAAGGCGCGGAGCTGAA 300
Qy 300 GCTCATAGCCAGGAGGCTCAACCCAGAGCTGGGTGTGAGTGTGAAGATGGTCTCAT 359
Db 301 GCTCATAGCCAGGAGGCTCAACCCAGAGCTGGGTGTGAGTGTGAAGATGGTCTCAT 360
Qy 360 GGGCAGGAGTGGAGCCGCAATCCAGAGGAGTGAAGCGCTGAAGAGGTTGAAG 419
Db 361 GGGCAGGAGTGGAGCCGCAATCCAGAGGAGTGAAGCGCTGAAGAGGTTGAAG 420
Qy 420 CAATCACAGTGTAAAAGAGACAGCTTGAATGATGACGCTGCTCTATGTTGGAAT 479
Db 421 CAATCACAGTGTAAAAGAGACAGCTTGAATGATGACGCTGCTCTATGTTGGAAT 480
Qy 480 TTGTTCAATTAATTTCTCCCAATAAGCTTTACAGCTTCTGCAAA 525
Db 481 TTGTTCAATTAATTTCTCCCAATAAGCTTTACAGCTTCTGCAAA 526

RESULT 10

AF058988

LOCUS

DEFINITION

ACCSSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

1..524

/organism="Homo sapiens"

/db_xref="taxon:9606"
/cell_line="LNCap; C4-2"
/note="isolated from prostate cancer cell lines by
differential display PCR; expression pattern correlates
with progression from androgen sensitive to androgen
insensitive in the human prostate cancer progression model
LNCap to C4-2"

80..433
/note="similar to GAGE family of melanoma antigens"

/codon_start=1

/product="melanoma antigen related GAGE-7"

/protein_id="AAC25989.1"

/db_xref="GI:3300090"

/translation="MSWRGRSTYYWPRRYVQPEVIGMPRPFQSFDEVEPATPEEG
EPATQRDPAAAGGEGDEGASAGOGPKPEADSBQGHQHPOTGCECEDGPDGQEVDPNP
EEVKTPEGEKQSQ"

BASE COUNT 146 a 113 c 152 g 111 t 2 others

ORIGIN

Query Match 92.5%; Score 497.6; DB 9; Length 524;
Best Local Similarity 97.9%; Pred. No. 3.9e-110;
Matches 513; Conservative 2; Mismatches 6; Indels 3; Gaps 1;

Qy 5 CAGGAGCTGTGAGGAGCTGCTGTGTCTCTCGCGTCCGGACTCTTTTCTCTACTG 64
Db 1 CAGGAGCTGTGAGGAGCTGCTGTGTCTCTCGCGTCCGGACTCTTTTCTCTACTG 60
Qy 65 AGATTTCATCTGTGTGAAATATGAGTTGGGAGGAGATGACCC--TATCGGCTTAGACC 121
Db 61 AGATTTCATCTGTGTGAAATATGAGTTGGGAGGAGATGACCTATTATTGGCTTAGACC 120
Qy 122 AAGAGCTACGTAGAGCTCTCTGAAATGATTGGGCTATCGGCGCGGAGCTTCAGTGA 181
Db 121 AAGAGCTACGTAGAGCTCTCTGAAATGATTGGGCTATCGGCGCGGAGCTTCAGTGA 180
Qy 182 TGAAGTGAACAGCAACACCTGAAGAGGGAACAGCAACTCAACGTCAGGATCCTGC 241
Db 181 TGAAGTGAACAGCAACACCTGAAGAGGGAACAGCAACTCAACGTCAGGATCCTGC 240
Qy 242 AGCTCTCAGGAGGAGGAGGATGAGGAGCATCTGCAGGTCAAGGCGCGGAGCTTGAAGC 301
Db 241 AGCTCTCAGGAGGAGGAGGATGAGGAGCATCTGCAGGTCAAGGCGCGGAGCTTGAAGC 300
Qy 302 TCATAGCCAGGAACAGGCTCACCCACAGACTGGGTGTGAGTGTGAAGATGGTCTCATGG 361
Db 301 TCATAGCCAGGAACAGGCTCACCCACAGACTGGGTGTGAGTGTGAAGATGGTCTCATGG 360
Qy 362 GCAGGAGATGGACCGCCCAATCCAGAGGAGTGAACCGCTGAAGAGAGGTGAAGAACCA 421
Db 361 GCAGGAGATGGACCGCCCAATCCAGAGGAGTGAACCGCTGAAGAGAGGTGAAGAACCA 420
Qy 422 ATCAGAGTGTAAAAGAGACACGTTGAAATGATGACGAGTGTCTCTATGTTGGAATTT 481
Db 421 ATCAGAGTGTAAAAGAGAGGACGTTGAAATGATGACGAGTGTCTCTATGTTGGAATTT 480
Qy 482 GTTCATTAATTTCTCCCAATAAGCTTTACAGCTTCTGCAAA 525
Db 481 GTTCATTAATTTCTCCCAATAAGCTTTACAGCTTCTGCAAA 524

RESULT 11

HSU19146

LOCUS

DEFINITION

ACCSSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

1 (bases 1 to 524)

/organism="Homo sapiens"

HSU19146
Human GAGE-5 protein mRNA, complete cds.

524 bp mRNA linear PRI 04-DEC-1995

U19146.1 GI:914906

human.

ORGANISM

REFERENCE

1 (bases 1 to 524)

/organism="Homo sapiens"

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS Van den Eynde,B., Peeters,O., De Backer,O., Gaugler,B., Lucas,S.
and Boon,T.
TITLE A new family of genes coding for an antigen recognized by
autologous cytolytic T lymphocytes on a human melanoma
J. Exp. Med. 182 (3), 689-698 (1995)
JOURNAL 95378788
MEDLINE
REFERENCE 2 (bases 1 to 524)
AUTHORS Van den Eynde,B.J.
Direct Submission
TITLE Submitted (28-DEC-1994) BENOIT J VAN DEN EYNDE, Ludwig Institute
JOURNAL for Cancer Research, 74 Avenue Hippocrate, Brussels B-1200, Belgium
FEATURES
source
1. .524
/organism="Homo sapiens"
/db_xref="taxon:9606"
/sex="female"
/cell_line="M2-MEL.43"
/tissue_type="melanoma"
/dev_stage="adult"
75. .428
/codon_start=1
/product="GAGE-5 protein"
/protein_id="AA82748.1"
/db_xref="GI:914907"
/translation="MSWRGRSTYYWPRRYVQPEVIGPMRPEQFSDEVEPATPEEG
EPATQRDPAAQEGEDEGASAGQPKPEADSQEQHPQTGCECEDGPDQEMDPNP
EEVKTPEGEKQSOC"
BASE COUNT 148 a 111 c 154 g 111 t
ORIGIN

Query Match 92.3%; Score 496.8; DB 9; Length 524;
Best Local Similarity 98.1%; Pred. No. 6.1e-110;
Matches 514; Conservative 0; Mismatches 7; Indels 3; Gaps 1;

Qy 10 AGCTGTAGGAGTCTGTGTGTTCTTCCTCGCTCGGACTCTTTTCCCTACTGAGATT 69
Db 1 AGCTGTAGGAGTCTGTGTGTTCTTCCTCGCTCGGACTCTTTTCCCTACTGAGATT 60
Qy 70 CATCTGTGTGAATATGAGTTGGCGAGGAAGATCGACC---TATCGGCCTAGACCAAGAC 126
Db 61 CATCTGTGTGAATATGAGTTGGCGAGGAAGATCGACCATTATTGGCCCTAGACCAAGGC 120
Qy 127 GCTAGCTAGAGCTCCTGAAATGATTGGCCCTATGCGGCCGAGCAGTTTCAGTGATGAAG 186
Db 121 GCTATGTACAGCTCCTGAGTGATTGGCCCTATGCGGCCGAGCAGTTTCAGTGATGAAG 180
Qy 187 TGGAAACAGCAACACTGGAAGAGGGGAACAGCAACTCAACGTTCAGGATCCTTCGACGTG 246
Db 181 TGGAAACAGCAACACTGGAAGAGGGGAACAGCAACTCAACGTTCAGGATCCTTCGACGTG 240
Qy 247 CTCAGAGGGAGAGGATGAGGAGCATCTGCAGGTCAAGGCCGAGCCCTGAAGCTCATA 305
Db 241 CTCAGAGGGAGAGGATGAGGAGCATCTGCAGGTCAAGGCCGAGCCCTGAAGCTGATA 300
Qy 307 GCCAGGAACAGGGTCAACCCACAGACTGGGTGTGAGTGTGAAGATGGTCTCTGATGGCAGG 366
Db 301 GCCAGGAACAGGGTCAACCCACAGACTGGGTGTGAGTGTGAAGATGGTCTCTGATGGCAGG 360
Qy 367 AGATGGACCCGCAATCAGAGGAGGTGAAAACGCTCAAGAGGTGAAAACCAATCAC 426
Db 361 AGATGGACCCGCAATCAGAGGAGGTGAAAACGCTCAAGAGGTGAAAACCAATCAC 420
Qy 427 AGTGTTAAAGAAGACACGTTGAAATGATGCAGGCTGCCTCTATGTTGGAAATTTGTTC 486
Db 421 AGTGTTAAAGAAGACACGTTGAAATGATGCAGGCTGCCTCTATGTTGGAAATTTGTTC 480
Qy 487 TTAATAATTCCTCCCAATAAAGCTTTACAGCCTTCTGCAAGAAAAA 530
Db 481 TTAATAATTCCTCCCAATAAAGCTTTACAGCCTTCTGCAAGAAAAA 524

RESULT 12

AF055474
LOCUS Homo sapiens GAGE-7B mRNA, complete cds.
DEFINITION AF055474
ACCESSION AF055474.1 GI:3511024
VERSION
KEYWORDS human,
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 526)
AUTHORS De Backer,O., Arden,K.C., Borettil,M., Vantomme,V., De Smet,C.,
Czekay,S., Viars,C.S., De plaen,E., Brasseur,F., Chomez,P., Van den
Eynde,B., Boon,T. and van der Bruggen,P.
TITLE Characterization of the GAGE genes that are expressed in various
human cancers and in normal testis
JOURNAL Cancer Res. 59 (13), 3157-3165 (1999)
MEDLINE 99323388
PUBMED 10397259
REFERENCE 2 (bases 1 to 526)
AUTHORS De Backer,O.R.Y.
TITLE Direct Submission
JOURNAL Submitted (24-MAR-1998) Ludwig Institute for Cancer Research,
Brussels Branch, 74 av. Hippocrate, Brussels B-1200, Belgium
FEATURES
source
1. .526
/organism="Homo sapiens"
/db_xref="taxon:9606"
55. .408
/codon_start=1
/product="GAGE-7B"
/protein_id="AAC33677.1"
/db_xref="GI:3511025"
/translation="MSWRGRSTYYWPRRYVQPEVIGPMRPEQFSDEVEPATPEEG
EPATQRDPAAQEGEDEGASAGQPKPEAHSEQHPQTGCECEDGPDQEMDPNP
EEVKTPEGEKQSOC"
BASE COUNT 168 a 109 c 143 g 106 t
ORIGIN

Query Match 90.7%; Score 488; DB 9; Length 526;
Best Local Similarity 98.4%; Pred. No. 8.2e-108;
Matches 504; Conservative 0; Mismatches 5; Indels 3; Gaps 1;

Qy 30 TGGTTCCTGCGCGTCCGGACTCTTTTCTCTACTGAGATTTCATCTGTGGAATATGAGT 89
Db 1 TGGTTCCTGCGCGTCCGGACTCTTTTCTCTACTGAGATTTCATCTGTGGAATATGAGT 60
Qy 90 TGGCGAGGAAGATCGACC---TATCGGCCTAGACCAAGACGCTACGTAGAGCCTCCTGAA 146
Db 61 TGGCGAGGAAGATCGACCTATTATTGGCTAGACCAAGCGCTATGTACAGCCTCCTGAA 120
Qy 147 ATGATTGGCCCTATCGCGCCGAGCAGTTTCAGTGATGAAGTGAACCAACACACCTGAA 206
Db 121 ATGATTGGCCCTATCGCGCCGAGCAGTTTCAGTGATGAAGTGAACCAACACACCTGAA 180
Qy 207 GAAGGGAAACAGCAACTCAACGTCAAGGATCCTGCAGTCTCAGAGGAGGAGAGTGAAG 266
Db 181 GAAGGGAAACAGCAACTCAACGTCAAGGATCCTGCAGTCTCAGAGGAGGAGAGTGAAG 240
Qy 267 GGAGCATCTGAGGTCAAGGCCGGAAGCCTGAAGCTCATAGCCAGGAACAGGGTCAACCA 326
Db 241 GGAGCATCTGAGGTCAAGGCCGGAAGCCTGAAGCTCATAGCCAGGAACAGGGTCAACCA 300
Qy 327 CAGACTGGGTGTGAGTGTGAAGATGGTCTCTGATGGCAGGAGATGAGCAGCCGCCAATCCA 386
Db 301 CAGACTGGGTGTGAGTGTGAAGATGGTCTCTGATGGCAGGAGATGAGCAGCCGCCAATCCA 360
Qy 387 GAGGAGGTGAACACCCCTGAAGAAGTGAAGAAGCAATCAGCTGTTAAAGAGACACAGT 446
Db 361 GAGGAGGTGAACACCCCTGAAGAAGTGAAGAAGCAATCAGCTGTTAAAGAGACAGT 420
Qy 447 TGAATGATGAGGCTGCTCCTATGTTGAAATTTTCTTCAATAAATTTCTCCCAATAAAG 506

|||||
Db 421 TGAATGATGAGGCTGCTCTATGTTGGAAATTTGTCATTAATAATTCCTCCCAATAAG 480
QY 507 CTTTACAGCCTTCTGCAAGAAAAA 538
Db 481 CTTTACAGCCTTCTGCAAGAAAAA 512
RESULT 13
AX284300/c 530 bp DNA linear PAT 20-NOV-2001
LOCUS Sequence 105 from Patent WO0179556.
DEFINITION AX284300
ACCESSION AX284300
VERSION AX284300.1 GI:17044988
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (sites)
AUTHORS Lillie,J., Brown,J.L., Bolt,A. and van Huffel,C.
TITLE Novel genes, compositions and methods for the identification,
assessment, prevention, and therapy of human cancers
JOURNAL Patent: WO 0179556-A 105 25-OCT-2001;
Millennium Predictive Medicine, Inc. (US)
FEATURES Location/Qualifiers
source 1..530
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 89 a 100 c 106 g 183 t 52 others
ORIGIN
Query Match 56.0%; Score 301.4; DB 6; Length 530;
Best Local Similarity 87.9%; Pred. No. 1.2e-62;
Matches 320; Conservative 0; Mismatches 42; Indels 2; Gaps 1;
QY 177 AGTGATGAACTGNAACGAC--AACACCTGAAGAGGGGAGGACCACTCAACGTGACG 234
Db 449 AGTNATNAAGTGACCCACNCAC--CCCTGAANAAGGNAACCACTCAACGTNAGG 390
QY 235 ATCCTGAGCTGTCTCAGGAGGAGGATGAGGAGGATCTGCAGGTCAAGGCCGAAGC 294
Db 389 ATCCTGNAGCTGCNAGNAGGATNAGGNGCATNTGCAGTCAAGGCCGAAGC 330
QY 295 CTGAAGCTATAGCCAGGACAGGTCACCCACAGACGTGGTGTGAGTGTGAAGATGGTC 354
Db 329 CTGAAGTGTAGCCAGGAACAGGTCNCCCCNACTGGGNGTNAGTGTNAATAATGGTC 270
QY 355 CTGATGGCAGGAGATGGACCGCCAAATCCAGAGGAGTGAACGCTCAAGAGAGTG 414
Db 269 CTNATGGCAGNATATGACCCNCAATCCAAAGNAGGTGAACNCCTGAANAAGGTG 210
QY 415 AAAGCAATCACAGTGTAAAGAACACAGTGTGAATGATGCAGGCTGCTCCTATGTTG 474
Db 209 AAAGCAATCCAGTGTAAANAAGGNCCTTGAATGATGCAGGCTGCTCCTATGTTG 150
QY 475 GAAATTTGTCATTAATAATTCCTCCCAATAAAGCTTTTACAGGCTTCTGCAAGAAAAA 534
Db 149 GAAATTTGTCATTAATAATTCCTCCCAATAAAGCTTTTACAGGCTTNTNAAAAAA 90
QY 535 AAAA 538
Db 89 AAAA 86
RESULT 14
AX284692/c 365 bp DNA linear PAT 20-NOV-2001
LOCUS Sequence 497 from Patent WO0179556.
DEFINITION AX284692
ACCESSION AX284692
VERSION AX284692.1 GI:17045380
KEYWORDS

SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (sites)
AUTHORS Lillie,J., Brown,J.L., Bolt,A. and van Huffel,C.
TITLE Novel genes, compositions and methods for the identification,
assessment, prevention, and therapy of human cancers
JOURNAL Patent: WO 0179556-A 497 25-OCT-2001;
Millennium Predictive Medicine, Inc. (US)
FEATURES Location/Qualifiers
source 1..365
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 63 a 86 c 66 g 150 t
ORIGIN
Query Match 55.6%; Score 299.2; DB 6; Length 365;
Best Local Similarity 95.9%; Pred. No. 4e-62;
Matches 307; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
QY 219 GCAACTCAACGCTCAGGATCCTGCAGCTGCTCAGAGGAGAGGATGAGGAGCATCTGCA 278
Db 365 GCAACTCAACGCTCAGGATCCTGCAGCTGCTCAGGAGGAAAGGATAAGGAGCATCTGCA 306
QY 279 GGTCAAGGCCGGAAGCCTGAAGCTCATAGCCAGGACACAGGTCACCCACAGACTGGGTGT 338
Db 305 GGTCAAGGCCGGAAGCCTGAAGCTCATAGCCAGGACACAGGTCACCCACAGACTGGGTGT 246
QY 339 GAGTGTGAAGATGGTCTCTGATGGGAGGAGATGGACCCCAATCCAGAGGAGGTGAAA 398
Db 245 GAGTGTGAAGATGGTCTCTGATGGGAGGAAATGGACCCCAATCCAAAGGAGGTGAAA 186
QY 399 ACGCTCAAGAGGTGAAAGCAATCAGACGTGTTAAAGACAGACGTTCAATGATGCA 458
Db 185 ACGCTCAAGAGGTGAAAGCAATCTCAGTGTAAAAAAGGACGTTGAAATGATGCA 126
QY 459 GGCTGCTCCTATGTGGAAATTTGTCATTAATAATTCCTCCCAATAAAGCTTTACAGCCTT 518
Db 125 GGCTGCTCCTATGTGGAAATTTGTCATTAATAATTCCTCCCAATAAAGCTTTACAGCCTT 66
QY 519 CTGCAAGAAAAA 538
Db 65 CTAAAAA 46
RESULT 15
HSA318881 493 bp mRNA linear PRI 14-JAN-2002
LOCUS Homo sapiens mRNA for XAGE-3 protein.
DEFINITION HSA318881
ACCESSION AJ318881
VERSION AJ318881.1 GI:18157211
KEYWORDS XAGE-3 gene.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (sites)
AUTHORS Zandman,A.J.W., van Kraats,A.A., Weidle,U.H., Ruiter,D.R. and Van
Muijen,G.N.P.
TITLE Expression profile of members of the XAGE cancer/testis antigen
family
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 493)
AUTHORS Zandman,A.J.W.
TITLE Direct Submission
JOURNAL Submitted (14-AUG-2001) Zandman A.J.W., Department of Pathology, PO
Box 9101 6500 HB Nijmegen, NETHERLANDS
FEATURES Location/Qualifiers
source 1..493
/organism="Homo sapiens"

```

/db_xref="taxon:9606"
/chromosome="X"
/map="Xp11.21-Xp11.22"
/tissue_type="Ewing sarcoma"
61..396
/gene="XAGE-3"
61..396
/gene="XAGE-3"
/codon_start=1
/product="XAGE-3 protein"
/protein_id="CAC83008.1"
/db_xref="GI:18157212"
/translation="MIWRKSTYRPRRSPVPPPELIGPMLPEPQDEEPQKPEPTESR
DPAPQEREDQAAETQVPLEADQLQSQTGGECNGPDDQGIILPKSEDFKMP
EGDROFQV"
BASE COUNT 167 a 98 c 125 g 103 t
ORIGIN

Query Match 33.9%; Score 182.4; DB 9; Length 493;
Best Local Similarity 68.9%; Pred. No. 7.3e-34;
Matches 317; Conservative 0; Mismatches 121; Indels 22; Gaps 4;

Qy 76 TGTGAATATGAGTTGGCGAGAGATCGACCTATCGCCTAGACCAAGACGCTACGTAG 135
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 53 TGTGAATATGATTTGGCGAGAGATCAACATATAGCCTAGGCCGAGGAGAAAGTGATC 112
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 136 AGCCTCTGAATGATTTGGGCTATGCGGCCCGAGCAGTTTCAGTGAAGTGGACCCAG 195
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 113 CACCTCTGAGCTGATTTGGGCTATGCGGCCCGAGCAGTTTCAGTGAAGTGGACCCAG 159
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 196 CAACACCTGAAGAGGGGAAACAGCACTCAACGTCAGGATCCTGCAGCTGCTCAGGAG- 254
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 160 -----CCTCAGCAAGAGGAAACCACTGAAAGTCGGATCCTGCACCTGCTCAGGAGA 214
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 255 --GGAGAGATCAGGAGCATCTGCAGTCAAGGCCGCGAAGCCTGNAAGCTCATAGCCAGG 312
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 215 GAGAAGAAGATCAGGGTCAGCTGAGACTCAAGTGCCTGACCTGGAAGCTGATCTCCAGG 274
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 313 AACAGGGTCACCCACAGACTGGGTGTGAGTGTGAAGATGGTCTGATGGCGCAGGAGATGG 372
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 275 AGCTGTCTCAGTCAAAAGACTGGGGGTGAATGTGGAATGGTCTGATGACCAAGGGGGA 334
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 373 ACCGCCCAATCCAGAGAGGTGAACGCCCTGAAGAGGTGAAGCAATCACAGTGT 432
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 335 TTCTGCCAAATCAGAACAAATTTAAATGCCAAGAGGAGTGACAGGCAACCAAGGTTT 394
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 433 AAAAGAAGACAGCTTGAATGATGACAGGCTGCTCCTATGTTGGAATTTGTTCAATAAAA 492
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 395 AAATGAGACAAAGCTGAACAAACACAAACTGT-TTTTATCTAAGATATTTGACTTAAA 453
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 493 TTCTCCCAATAAGCTTTACAGCTTCTGCAAGAAAAA 532
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 454 ATATCGAAATAAATTTTTCAGCTTTCTCCAAAAA 493
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

Search completed: July 1, 2002, 07:41:59
Job time: 3730 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 1, 2002, 06:37:34 ; Search time 1595.88 seconds
(without alignments)
4550.068 Million cell updates/sec

Title: US-09-782-745-14
Perfect score: 538
Sequence: 1 ACGCCAGGAGCTGTGAGC.....CTGCCAAGAAAAA 538

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues
Total number of hits satisfying chosen parameters: 17277834

Minimum DB seq length: 0
Maximum DB seq length: 538

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST.*
1: em_estba.*
2: em_esthum.*
3: em_estin.*
4: em_estmu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_hic.*
9: gb_est1.*
10: gb_est2.*
11: gb_hic.*
12: gb_gss.*
13: em_gss_hum.*
14: em_gss_inv.*
15: em_gss_pln.*
16: em_gss_vrt.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	492	91.4	527	10	BI826605	BI826605 603077056
2	476.8	88.6	509	10	BI868671	BI868671 603392594
3	452.4	84.1	489	9	AA447559	AA447559 zw81e11.s
4	431.4	80.2	457	9	AW510753	AW510753 hd39d05.x
5	424.6	78.9	447	9	AI381509	AI381509 te76b07.x
6	424.4	78.9	464	9	AW738037	AW738037 nx15e11.s
7	413.8	76.9	418	9	AW016546	AW016546 UT-H-B10p
8	413.4	76.8	480	10	BG120336	BG120336 602353732
9	412.8	76.7	426	9	AA868226	AA868226 ak48h07.s
10	410.8	76.4	450	9	AA760996	AA760996 nx32h08.s
11	401.2	74.6	455	9	AI187350	AI187350 qf29a05.x
12	365.4	67.9	412	9	AW102587	AW102587 xd67g07.x
13	360.4	67.0	445	11	BC005363	BC005363 Homo sapi
14	359.4	66.8	419	9	AA448542	AA448542 zw81e11.r
15	329.8	61.3	384	9	AA913206	AA913206 om53f03.s
16	318.8	59.3	397	10	BG206349	BG206349 RST25797
17	236.4	55.1	398	9	AA918604	AA918604 ol53f02.s

C	18	250.4	46.5	256	10	BG181480
	19	233.4	43.4	258	10	BG186708
	20	228.2	42.4	253	10	BG184057
	21	226.6	42.1	261	10	BG199060
	22	224.8	41.8	245	10	BG212621
	23	219.8	40.9	292	10	BG212622
	24	214.8	39.9	245	10	BG208433
	25	213.4	39.7	524	11	AF318372
	26	207.6	38.6	507	9	AA972716
	27	207	38.5	256	10	BG216461
	28	204.8	38.1	224	9	AA738394
	29	202.4	37.6	382	10	BF869799
	30	200.6	37.3	499	10	BG354572
	31	193.8	36.0	197	9	AI968311
	32	192.8	35.8	226	10	BG220441
	33	186.8	34.7	503	9	AI742551
	34	183.8	34.2	468	9	AI128458
	35	181.8	33.8	453	9	AI189715
	36	178.4	33.2	483	9	AI149003
	37	177.8	33.0	496	9	AI188535
	38	176.8	32.9	468	9	AA039331
	39	174.4	32.4	496	9	AA993199
	40	173.6	32.3	280	10	X93834
	41	172.4	32.0	445	10	BE302172
	42	171	31.8	532	9	AA737311
	43	170.4	31.7	429	9	AI189067
	44	170	31.6	437	9	AI740470
	45	169.6	31.5	394	10	D59091

ALIGNMENTS

RESULT 1
BI826605
LOCUS 603077056F1 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5168892 5',
DEFINITION mRNA sequence.
ACCESSION BI826605
VERSION BI826605.1 GI:15938155
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 527)
NIH-MGC http://mgc.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strassberg, Ph.D.
COMMENT Email: cgapbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLNL119 row: p column: 13
High quality sequence stop: 519.
FEATURES
Location/Qualifiers
1..527
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5168892"
/clone_lib="NIH_MGC_119"
/tissue_type="medulla"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMV-Sport6; Site: 1: NotI;
Site 2: EcoRV (destroyed); RNA source normal medulla from
anonymous male age 27. Library is oligo-dT primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.3 kb, insert size range

0.9-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 013. Note: this is a NIH_MGC Library."

BASE COUNT 154 a 111 c 151 g 111 t
ORIGIN

Query'Match 91.4%; Score 492; DB 10; Length 527;
Best Local Similarity 98.3%; Pred. No. 3.8e-104;
Matches 519; Conservative 0; Mismatches 5; Indels 4; Gaps 2;

Qy 12 CTGTGAGGAGCTGTGTGGTTCCTGCGCGTCCGGAGCTCTTTTCCCTACTGAGATTCA 71
Db 1 CTGTGAGGAGCTGTGTGGTTCCTGCGCGTCCGGAGCTCTTTTCCCTACTGAGATTCA 60
Qy 72 TCTGTGTGAAATATGAGTTGGCGAGGAAGATCGACC---TATCGGCCTAGACCAAGAGCGC 128
Db 61 TCTGTGTGAAATATGAGTTGGCGAGGAAGATCGACCTATTATTGGCCTAGACCAAGGCGC 120
Qy 129 TAGGTAGAGCCCTCTGAAATGATGGGCCCTATGCGGCCCGAGCAGTTCAGTGAAGTG 188
Db 121 TAGGTAGAGCCCTCTGAAATGATGGGCCCTATGCGGCCCGAGCAGTTCAGTGAAGTG 180
Qy 189 GAACAGCAACACCTCTGAAGAAAGGGGAACCAAGCAACTCAAGTCAAGGATCCTGCAAGTCT 248
Db 181 GAACAGCAACACCTCTGAAGAAAGGGGAACCAAGCAACTCAAGTCAAGGATCCTGCAAGTCT 240
Qy 249 CAGGAGGAGAGATGAGGAGCATCTGCAGGTCAAGGGCGGAAGCCTGAAGCTCATAGC 308
Db 241 CAGGAGGAGAGATGAGGAGCATCTGCAGGTCAAGGGCGGAAGCCTGAAGCTCATAGC 300
Qy 309 CAGGAACAGGGTCACCCACAGACTGGGTGTGAGTGTGAAGATGTCCTGATGGGCGAGG 368
Db 301 CAGGAACAGGGTCACCCACAGACTGGGTGTGAGTGTGAAGATGTCCTGATGGGCGAGG 360
Qy 369 ATGGACCCGCCAATCCAGAGGAGTGAAACCGCTCAAGAAAGTGAAACCAATCACAG 428
Db 361 ATGGACCCGCCAATCCAGAGGAGTGAAACCGCTCAAGAAAGTGAAACCAATCACAG 420
Qy 429 TGTAAAAGAAGACAGCTTTGAAATGATGCAGGCTGCTCTATGTTGAAATTTGTTCATT 488
Db 421 TGTAAAAGAAGACAGCTTTGAAATGATGCAGGCTGCTCTATGTTGAAATTTGTTCATT 480
Qy 489 AAAATCTCCCAATAGCTTTACAGCTTCTCAAGAAAGAAAAA 536
Db 481 AAAATCTCCCAATAGCTTTACAG--CTTCTCAAGAAAGAAAAA 527

RESULT 2
BI868671
LOCUS 603392594F1 NIH_MGC_90 Homo sapiens cDNA clone IMAGE:5402663 5',
DEFINITION mRNA sequence.
ACCESSION BI868671
VERSION BI868671.1 GI:16042344
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:

http://image.llnl.gov
Plate: LLAM12026 Row: 1 Column: 24
High quality sequence stop: 509.
Location/Qualifiers
1. 509

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5402663"
/issue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: liver; Vector: pCMV-Sport6; Site: 1; NotI;
Site: 2; SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.7 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

BASE COUNT 142 a 107 c 154 g 106 t
ORIGIN

Query'Match 88.6%; Score 476.8; DB 10; Length 509;
Best Local Similarity 98.0%; Pred. No. 1.3e-100;
Matches 494; Conservative 0; Mismatches 7; Indels 3; Gaps 1;

Qy 1 ACCCCAGGAGCTGTGAGGAGCTGTGTGGTTCCTGCGGCCGAGCTCTTTTCTCTCT 60
Db 6 ACCTAGGAGCTGTGAGGAGCTGTGTGGTTCCTGCGGCCGAGCTCTTTTCTCTCT 65
Qy 61 ACTGAGATTCATCTGTGAAATATGATTTGGCGAGGAGATCGAC---TATCGGCCCTA 117
Db 66 ACTGAGATTCATCTGTGAAATATGATTTGGCGAGGAGATCGACCTATTATTGGCCCTA 125
Qy 118 GACCAAGACCTCTAGAGCTCTCTGAAATGATTTGGCCCTATGCGGCCGAGCTCTCA 177
Db 126 GACCAAGGCGCTATGTACAGCTCTCTGAAATGATTTGGCCCTATGCGGCCGAGCTCTCA 185
Qy 178 GTGATGAGTGGACCAAGCAACCTGAGAGGAGGAGCACTCAACCTCAACCTCAAGTTC 237
Db 186 GTGATGAGTGGACCAAGCAACCTGAGAGGAGGAGCACTCAACCTCAACCTCAAGTTC 245
Qy 238 CTGACGCTGCTCAGGAGGAGGAGGATGAGGAGCATCTGCAGGTCAAGGCGCGAAGCCTG 297
Db 246 CTGACGCTGCTCAGGAGGAGGAGGATGAGGAGCATCTGCAGGTCAAGGCGCTGAAGCCTG 305
Qy 298 AGCTCATAGCCAGGACAGGTCACCCACAGCTGCGGTGTGAGTGTGAAGATGTCCTG 357
Db 306 AGCTCATAGCCAGGACAGGTCACCCACAGCTGCGGTGTGAGTGTGAAGATGTCCTG 365
Qy 358 ATGGGAGGAGATGGACCCGCCAAATCCAGAGGAGTGAACACGCTGAAGAGGTGAAA 417
Db 366 ATGGGAGGAGATGGACCCGCCAAATCCAGAGGAGTGAACACGCTGAAGAGGTGAAA 425
Qy 418 AGCAATCACAGTGTAAAAGAGACACGTTGAATGATGAGGCTGCTCTATGTTGGAA 477
Db 426 AGCAATCACAGTGTAAAAGAGACGCTGAAATGATGAGGCTGCTCTATGTTGGAA 485
Qy 478 ATTGTTCATTAAATTTCTCCAA 501
Db 486 ATTGTTCATTAAATTTCTCCAA 509

RESULT 3
AA447559/c
LOCUS
DEFINITION 3' similar to Tr:G914905 G914905 GAGE-4 PROTEIN. [1] ;, mRNA
sequence.
ACCESSION AA447559
VERSION AA447559.1 GI:2161229
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1 (bases 1 to 489)
AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theising, B., White, Y., Wyllie, T., Waterston, R., and Wilson, R.
TITLE WashU-Merck EST Project 1997
JOURNAL Unpublished (1997)
COMMENT Contact: Wilton RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -41m13 fwd. Et from Amersham
High quality sequence stop: 427.

FEATURES
source
1..489
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:782636"
/clone_lib="Soares_testis_NHT"
/sex="male"
/lab_host="DH10B"
/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from Clontech Laboratories , Inc., and primed with a Not I - oligo(dT) primer [5', TGTTACCAATCGAATGGAGCGGCCCAATTTTTTTTTTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization to Cot5, and was constructed by Bento Soares and M. Fatima Bonaldo."
101 a 139 c 105 g 144 t

BASE COUNT
ORIGIN
Query Match 84.1%; Score 452.4; DB 9; Length 489;
Best Local Similarity 98.0%; Pred. No. 6e-95;
Matches 480; Conservative 0; Mismatches 6; Indels 4; Gaps 2;

Qy 43 CCGGACTCTTTTCTCTACTGAGATTTCATCTGTGTGAAATATGAGTTGGCGAGGAAGAT 102
Db 489 CCGGACTCTTTTCTCTACTGAGATTTCATCTGTGTGAAATATGAGTTGGCGAGGAAGAT 430
Qy 103 CGACC---TATCGGCTAGACCCAGACGCTACGTAGAGCCTCTGAAATGATTGGCCCTA 159
Db 429 CGACCTATTATTGGCCTAGACCAAGCGCTATGTACAGCCTCCTGAAGTGATT-GGCCCTA 371
Qy 160 TCGGCGCCGAGCAGTTCACTGATGAAGTGAAGCAACACCTGAAGAGGGGAACCCAG 219
Db 370 TCGGCGCCGAGCAGTTCACTGATGAAGTGAAGCAACACCTGAAGAGGGGAACCCAG 311
Qy 220 CAACTCAACGTCAGGATTCCTGACGTGCTCAGGAGGAGAGGATGAGGAGCATCTGCAG 279
Db 310 CAACTCAACGTCAGGATTCCTGACGTGCTCAGGAGGAGAGGATGAGGAGCATCTGCAG 251
Qy 280 GTCAGGGCCGAGCCTGAAGCTCATGCCAGGAGAGGGTCAACCCAGACTGGGTGG 339
Db 250 GTCAGGGCCGAGCCTGAAGCTCATGCCAGGAGAGGGTCAACCCAGACTGGGTGG 191
Qy 340 AGTGTGAAGATGGTCTCTGATGGCAGGAGATGGACCCGCAAAATCCAGAGGAGGTGA 399
Db 190 AGTGTGAAGATGGTCTCTGATGGCAGGAGATGGACCCGCAAAATCCAGAGGAGGTGA 131
Qy 400 CGCCTGAAGAAGGTGAAGAATACAGTGTAAAAAGAACACAGCTTTGAATGATGCAG 459
Db 130 CGCCTGAAGAAGGTGAAGAATACAGTGTAAAAAGAACACAGCTTTGAATGATGCAG 71
Qy 460 GCTGCTCCTATGTTGGAAATTTGTTCAATTAATAATCTCCCAATAAAGCTTTACAGCCTTC 519

Db 70 GCTCCTCTATGTTGGAAATTTGTTCAATAAAATTCCTCCCAATAAGCTTTACAGCCCTTC 11
Qy 520 TGCAAAAGAAA 529
Db 10 TGCAAAAGAAA 1

RESULT 4
AW510753/c
LOCUS
DEFINITION
hd39d05.x1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone
IMAGE:2911881 3' similar to SW:GGE4_HUMAN Q13068 GAGE-4 PROTEIN. ;
mRNA sequence.
ACCESSION
AW510753
VERSION
AW510753.1 GI:7148831
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 457)
REFERENCE
AUTHORS
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL
Unpublished (1997)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40UP from Gibco.

FEATURES
source
1..457
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2911881"
/clone_lib="Soares_NFL_T_GBC_S1"
/lab_host="DH10B"
/note="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; Equal amounts of plasmid DNA from three normalized libraries (fetal lung NbHL19W, testis NHT, and B-cell NCI-CGAP-GCB1) were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of I.M.A.G.E. clones 297480-302087, 682632-687239, 726408-728711, and 729096-731399. Subtraction by Bento Soares and M. Fatima Bonaldo."
87 a 133 c 95 g 142 t

BASE COUNT
ORIGIN
Query Match 80.2%; Score 431.4; DB 9; Length 457;
Best Local Similarity 98.0%; Pred. No. 4.5e-90;
Matches 448; Conservative 0; Mismatches 6; Indels 3; Gaps 1;

Qy 77 GTGAATATGAGTTGGCGAGGAAGATCGACATATC---GGCCTAGACCAAGACGCTACGT 133
Db 457 GTGAATATGAGTTGGCGAGGAAGATCGACATATATTTGGCCTAGACCAAGCGCTATGT 398
Qy 134 AGAGCCTCTCTGAATGATTTGGGCTATTCGGCCCGCAGGAGTTTCAGTGATGAAGTGAACC 193
Db 397 ACAGCCTCTCTGAATGATTTGGGCTATTCGGCCCGCAGGAGTTTCAGTGATGAAGTGAACC 338
Qy 194 AGCAACACCTGAGAAGGGGAACCCAGCAACTCAAGCTCAGGATCTCTGACGTCTCTCAGGA 253
Db 337 AGCAACACCTGAGAAGGGGAACCCAGCAACTCAAGCTCAGGATCTCTGACGTCTCTCAGGA 278
Qy 254 GGGAGGAGGATGAGGGAGCATCTGCAGGTTCGAAGGGCCGAAGCCTGAAGCTCATAGCCAGGA 313
Db 277 GGGAGGAGGATGAGGGAGCATCTGCAGGTTCGAAGGGCCGAAGCCTGAAGCTCATAGCCAGGA 218

```
QY 314 ACAGGTCACCCACAGACTGGGTGTGAGTGTGAAGATGGTCCTGATGGCGAGGAGATGGA 373
|||||
Db 217 ACAGGGTCACCCACAGACTGGGTGTGAGTGTGAAGATGGTCCTGATGGCGAGGAGATGGA 158
|||||
QY 374 CCGGCCAAATCCAGAGGAGGTGAACCGCTCGAAGAGGTGAAGCAATCACAGTGTTA 433
|||||
Db 157 CCGGCCAAATCCAGAGGAGGTGAACCGCTCGAAGAGGTGAAGCAATCACAGTGTTA 98
|||||
QY 434 AAAGAAGACAGCTGAATGATGACGGCTGCTCCTATGTTGGAATTTGTCATTAAT 493
|||||
Db 97 AAAGAAGCAGCTGAATGATGACGGCTGCTCCTATGTTGGAATTTGTCATTAAT 38
|||||
QY 494 TCTCCCAATAAGCTTTACAGCCTTCTGCAAAAGAAA 530
|||||
Db 37 TCTCCCAATAAGCTTTACAGCCTTCTGCAAAAGAAA 1
|||||

RESULT 5
AI381509/c
LOCUS
DEFINITION
IMAGE:2092597.3' similar to TR:Q13070 Q13070 GAGE-6 PROTEIN. ;
mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 447)
NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
JOURNAL
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-f@mail.nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 497 Std Error: 0.00
Seq primer: -40UP from Gibco.
FEATURES
Location/Qualifiers
1..447
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="IMAGE:2092597"
/clone_lib="Soares_NFL_T_GBC_S1"
/lab_host="DH10B"
/note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
a modified polylinker; Site_1: Not I; Site_2: Eco RI;
Equal amounts of plasmid DNA from three normalized
libraries (fetal lung NBHL19W, testis NHT, and B-cell
NCI-CCGAP.GCBI) were mixed, and ss circles were made in
vitro. Following HAP purification, this DNA was used as
tracer in a subtractive hybridization reaction. The driver
was PCR-amplified cDNAs from pools of 5,000 clones made
from the same 3 libraries. The pools consisted of
I.M.A.G.E. clones 297480-302087, 682632-687239,
726408-728711, and 729096-731399. Subtraction by Bento
Soares and M. Fatima Bonaldo."
84 a 130 c 95 g 138 t

Query Match 78.9%; Score 424.6; DB 9; Length 447;
Best Local Similarity 98.4%; Pred. No. 1.7e-88;
Matches 440; Conservative 0; Mismatches 4; Indels 3; Gaps 1;

QY 88 GTTGGCGGAGAGATGCACGAC---TATCGGCTAGACCAAGCGCTACGTACAGCCTCCCTG 144
|||||
Db 447 GTTGGCGGAGAGATGCACCTATTATTGGCTAGACCAAGCGCTATGTACAGCCTCCCTG 388
|||||
```

```
QY 145 AAATGATTGGCCCTATGCGGCCCGAGCACCTTCAGTGATGAAGTGAACCCAGCAACACCTG 204
|||||
Db 387 AAATGATTGGCCCTATGCGGCCCGAGCACCTTCAGTGATGAAGTGAACCCAGCAACACCTG 328
|||||
QY 205 AAGAAGGGGAACCCAGCAACTCAACGTCAGGATCCTCGAGTCAGTGTCTCAGGAGGAGAGGATG 264
|||||
Db 327 AAGAAGGGGAACCCAGCAACTCAACGTCAGGATCCTCGAGTCAGTGTCTCAGGAGGAGAGGATG 268
|||||
QY 265 AGGAGCATCTGCAAGGTCAAGGCCGGAAGCCTGAAGCTCTATAGCCAGGAACAGGGTCAACC 324
|||||
Db 267 AGGAGCATCTGCAAGGTCAAGGCCGGAAGCCTGAAGCTCTATAGCCAGGAACAGGGTCAACC 208
|||||
QY 325 CACAGACTGGGTGTGAGTGTGAAGATGGTCCTGATGGCAGGAGATGAGCCGCCAATC 384
|||||
Db 207 CACAGACTGGGTGTGAGTGTGAAGATGGTCCTGATGGCAGGAGATGAGCCGCCAATC 148
|||||
QY 385 CAGAGGAGGTGAACCGCTGAAGAGGTGAAGCAANTCACAGTGTAAAGAGAGACAC 444
|||||
Db 147 CAGAGGAGGTGAACCGCTGAAGAGGTGAAGCAANTCACAGTGTAAAGAGAGACAC 88
|||||
QY 445 GTTGAATGATGCAAGGCTGCTCCTATGTTGGAATTTGTCATTAATTTCTCCCAATAA 504
|||||
Db 87 GTTGAATGATGCAAGGCTGCTCCTATGTTGGAATTTGTCATTAATTTCTCCCAATAA 28
|||||
QY 505 AGCTTTACAGCCTTCTGCAAAAGAAA 531
|||||
Db 27 AGCTTTACAGCCTTCTGCAAAAGAAA 1
|||||

RESULT 6
AA738037/c
LOCUS
DEFINITION
IMAGE:464 bp mRNA linear EST 22-JAN-1998
similar to TR:Q13070 Q13070 GAGE-6 PROTEIN. ; mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 464)
NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
JOURNAL
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-f@mail.nih.gov
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CCGAP clone distribution Information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 511 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 435.
FEATURES
Location/Qualifiers
1..464
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="IMAGE:1256204"
/clone_lib="NCI-CCGAP_GC3"
/tissue_type="pooled germ cell tumors"
/lab_host="DH10B"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; 1st strand cDNA was prepared from 3 pooled
germ cell tumors, and was then primed with a Not I -
oligo(dT) primer. Double-stranded cDNA was ligated to Eco
RI adaptors (Pharmacia), digested with Not I and cloned
into the Not I and Eco RI sites of the modified pT7T3
```

```
vector. Library is not normalized. Library was
constructed by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT      86 a      132 c      93 g      153 t
ORIGIN

Query Match      78.9%; Score 424.4; DB 9; Length 464;
Best Local Similarity 96.9%; Pred. NO. 1.9e-88;
Matches 444; Conservative 0; Mismatches 11; Indels 3; Gaps 1;

QY 84 ATGAGTTGGCAGGAAGATCGACC---TATCGGCCTAGACCAGAGAGCGCTACGTAGAGCCT 140
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 464 ATGAGTTGGCAGGAAGATCGACCTGTTATTGGCCTAGTCCAGAGCGCTATGTACAGCCT 405
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 141 CCTCAAAATGATTGGCCCTATGCGGCCCGAGCAGTTCAGTGTGAAGTGGAAACAGCAACA 200
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 404 CCTCAAAATGATTGGCCCTATGCGGCCCGAGCAGTTCAGTGTGAAGTGGAAACAGCAACA 345
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 201 CCTGAAGAAGGGGAACCAAGCAACTCAACGTCAGGATCCTGCAGCTGCTCAGGAGGGAGAG 260
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 344 CCTGAAGAAGGGGAACCAAGCAACTCAACGTCAGGATCCTGCAGCTGCTCAGGAGGGAGAG 285
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 261 GATCAGGGAGCATCTGCAGGTCAAGGCCGAAGCCTGAAGCTCATAGCCAGGAACAGGGT 320
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 284 GATCAGGGAGCATCTGCAGGTCAAGGCCGAAGCCTGAAGCTCATAGCCAGGAACAGGGT 225
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 321 CACCACAGACTGGGTGTGAGTGTGAAGATGCTCTGATGGCGAGGAGATGGACCCGCCA 380
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 224 CACCACAGACTGGGTGTGAGTGTGAAGATGCTCTGATGGCGAGGAGATGGACCCGCCA 165
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 381 AATCCAGAGGAGTGAACCGCTGAAGAAGTGAAGAGCAATCACAGTGTAAAAAGAAG 440
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 164 AATCCAGAGGAGTGAACCGCTGAAGAAGTGAAGAGCAATCACAGTGTAAAAAGAAG 105
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 441 ACAGCTTCAAAATGATGCAGGCTGCTCTATGTTGGAAATTTGTTCAATTAATAATCTCCCA 500
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 104 GCAGCTTCAAAATGATGCAGGCTGCTCTATGTTGGAAATTTGTTCAATTAATAATCTCCCA 45
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 501 ATAAAGCTTTACAGCTTCTGCAAGAAGAAAAA 538
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 44 ATAGAGTTTACAGCTTCTGCAAAA 7
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 7
AW016546/c
LOCUS
DEFINITION
  UI-H-BiOp-abg-g-06-0-UI.sl NCI_CGAP_Sub2 Homo sapiens cdNA clone
  IMAGE:2711986 3', mRNA sequence.
ACCESSION
  AW016546
VERSION
  AW016546.1 GI:5865303
KEYWORDS
  EST.
SOURCE
  human.
ORGANISM
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1 (bases 1 to 418)
  NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
  Tumor Gene Index
  Unpublished (1997)
  Contact: Robert Strausberg, Ph.D.
  Email: cgapbs-r@mail.nih.gov
  The sequence contained an oligo-dT track that was present in the
  oligonucleotide that was used to prime the synthesis of first
  strand cDNA and therefore this may represent a bonafide poly A
  tail. cDNA Library Preparation: M.B. Soares Lab Clone distribution:
  NCI-CGAP clone distribution information can be found through the
  I.M.A.G.E. Consortium/LINL at:
  www-bio.lnl.gov/bbrp/image/image.html
  Seq primer: M13 Forward
  POLYA=Yes.
FEATURES
  source
  1..418
```

```
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2711986"
/clone_lib="NCI_CGAP_Sub2"
/lab_host="PH10B (Life Technologies)"
/note="Vector: pT73D-pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; The
NCI_CGAP_Sub2 library is a subtracted library derived from
BI. BI constitutes a mixture of 21 normalized or
subtracted NCI_CGAP libraries: NCI_CGAP_Co4, NCI_CGAP_Pr22
, NCI_CGAP_Pr28, NCI_CGAP_Co10, NCI_CGAP_Co16,
NCI_CGAP_Kid5, NCI_CGAP_Kid12, NCI_CGAP_Kid3,
NCI_CGAP_Kid11, NCI_CGAP_Lym2, NCI_CGAP_Br2, NCI_CGAP_Co8
, NCI_CGAP_LL1, NCI_CGAP_Le12, NCI_CGAP_Brn23, NCI_CGAP_Gc5
, NCI_CGAP_Lu24, NCI_CGAP_Lu19, NCI_CGAP_Gc4, NCI_CGAP_Gc6
, NCI_CGAP_Brn25. These 21 libraries were pooled and a
single-stranded DNA preparation of the resulting mixture
was used as a tracer in a subtractive hybridization with a
driver whose composition is detailed below: NCI_CGAP_Kid3
pool 1 LHAM 3334-3337, 3682-3683, 3798-3803 (IMAGE)
CloneIDs 1322376-1323911, 1456008-1456775, 1500552-1502855
NCI_CGAP_Kid5 pool 1 LHAM 3338-3342, 3722-3725,
3776-3778 (IMAGE) CloneIDs 1323912-1325831,
1471368-1472903, 1492104-1493255) NCI_CGAP_Lu5 pool 1 LHAM
3575-3582, 3851-3854 (IMAGE) CloneIDs 1414920-1417991,
1520904-1522439) NCI_CGAP_Gc4 pool 1 LHAM 3164-3167,
3716-3720, 3733-3735 (IMAGE) CloneIDs 1257096-1258631,
1469064-1470983, 1475592-1476743) NCI_CGAP_Pr22 pool 1
LHAM 2457-2459, 2758-2759, 3062-3068 (IMAGE) CloneIDs
985608-986759, 1101192-1101959, 1217928-1220615)
NCI_CGAP_Co10 pool 1 LHAM 2644-2653, 2871-2872 (IMAGE)
CloneIDs 1057416-1061255, 1144584-1145351) The resulting
subtracted library contained 4 million recombinants.
Subtraction was performed as previously described [Bonaldo
, Lennon & Soares (1996): Normalization and Subtraction:
Two Approaches To Facilitate Gene Discovery. Genome
Research 6, 791-806.
TAG_LIB=NCI_CGAP_Gc4
TAG_TISSUE=germ cell
TAG_SEQ=AAATC"

BASE COUNT      74 a      119 c      87 g      138 t
ORIGIN

Query Match      76.9%; Score 413.8; DB 9; Length 418;
Best Local Similarity 99.5%; Pred. NO. 5.5e-86;
Matches 415; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 122 AAGACGCTACGTAGAGCCTCCTGAAATGATTGGGCTTATGCGGCCCGAGCAGTTCAGTGA 181
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 418 AAGACGCTACGTAGAGCCTCCTGAAATGATTGGGCTTATGCGGCCCGAGCAGTTCAGTGA 359
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 182 TGAAGTGGAAACCAACACACCTGAAGAAGGGGAACCACTCAACGTACAGATCCTGC 241
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 358 TGAAGTGGAAACCAACACACCTGAAGAAGGGGAACCACTCAACGTACAGATCCTGC 299
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 242 AGCTGCTCAGGAGGAGGATGAGGGAGCATCTCAGGTCAAGGGCCGAAGCCTGAAGC 301
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 298 AGCTGCTCAGGAGGAGGATGAGGGAGCATCTCAGGTCAAGGGCCGAAGCCTGAAGC 239
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 302 TCATAGCCAGGAACAGGGTACCCACAGACTGGGTGTGAGTGTGAAGATGGTCTGATGG 361
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 238 TCATAGCCAGGAACAGGGTACCCACAGACTGGGTGTGAGTGTGAAGATGGTCTGATGG 179
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 362 GCAGGAGATGACCCGCCAATCCAGAGGAGTGAACCGCTCGAGAGAGGTGAAGAAGCA 421
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 178 GCAGGAGATGACCCGCCAATCCAGAGGAGTGAACCGCTCGAGAGAGGTGAAGAAGCA 119
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 422 ATCACAGTGTAAAGAAGACACAGTGAATGATCAGGCTGCTCCTATGTTGGAATTT 481
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 118 ATCACAGTGTAAAGAAGAGGCAGTGAATGATCAGGCTGCTCCTATGTTGGAATTT 59
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 482 GTTCATTAATAATCTCCCAATAAAGCTTTACAGCCTTCTGTGCAAAAGAAAAA 538
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
Db 58 GTTCATTAAATCTCCCAATAAAGCTTTACAGCCTTCTGCAAAAAA 2
|||||
RESULT 8
BG120336
LOCUS
DEFINITION 602353732F1 NTH_MGC_90 480 bp mRNA linear EST 30-JAN-2001
mRNA sequence.
ACCESSION BG120336
VERSION BG120336.1 GI:12613845
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 480)
AUTHORS NTH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM10239 row: p column: 07
High quality sequence stop: 480.
Location/Qualifiers
1. 480
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4451958"
/clone_lib="NTH_MGC_90"
/tissue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: liver; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.7 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NTH_MGC library."
BASE COUNT 170 a 93 c 133 g 84 t
ORIGIN
source
Query Match 76.8%; Score 413.4; DB 10; Length 480;
Best Local Similarity 97.4%; Pred. No. 6.8e-86;
Matches 420; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
QY 108 TATCGGCTACGACCAAGACGCTACGTAGAGCCTCTCGAATGATTGGCCCTATGCGGCC 167
|||
Db 15 TATTGGCTAGACCAAGCGCTATGTACAGCCTCTCTGAAGTGTATTGGCCTATGCGGCC 74
|||||
QY 168 GAGCAGTTCAGTGAAGTGGACCCAGCAACACCTCGAAGAGGGGAACAGCAACTCAA 227
|||||
Db 75 GAGCAGTTCAGTGAAGTGGAAACCCAGCAACACCTCGAAGAGGGGAACAGCAACTCAA 134
|||||
QY 228 CGTCAGGATCCTGCAGCTGCTCAGAGGAGGAGGATGAGGAGCATCTGCAGGTCAGGG 287
|||||
Db 135 CGTCAGGATCCTGCAGCTGCTCAGAGGAGGAGGATGAGGAGCATCTGCAGGTCAGGG 194
|||||
QY 288 CCGAAGCCTGAAGCTCATAGCCAGGAACAGGGTCAACCCACAGACTGGGTGTAGTGTAA 347
|||||
Db 195 CCGAAGCCTGAAGCTGATAGCCAGGAACAGGGTCAACCCACAGACTGGGTGTAGTGTAA 254
|||||
QY 348 GATGTCCTGATGGCAGGAGATGGACCCGCCAATCCAGAGGAGGTGAAACGCCTGAA 407
|||||
Db 255 GATGTCCTGATGGCAGGAGATGGACCCGCCAATCCAGAGGAGGTGAAACGCCTGAA 314
|||||
QY 408 GAAGGTGAAAGCAATCACAGTGTATAAAGAACACACCTGTTGAATGATCGAGCTGCTCC 467
|||||
```

```
Db 315 GAAGGTGAAAGCAATCACAGTGTATAAAGAGGACAGCTTGAATGATCGAGCTGCTCC 374
|||||
QY 468 TATGTTGGAATTTGTTCAATAAATCTCCCAATAAAGCTTTACAGCCTTCTGCAAGA 527
|||||
Db 375 TATGTTGGAATTTGTTCAATAAATCTCCCAATAAAGCTTTACAGCCTTCTGCAAAA 434
|||||
QY 528 AAAAAA 538
|||||
Db 435 AAAAAA 445
|||||
RESULT 9
AA868226/c
LOCUS
DEFINITION ak48h07.s1 Soares_testis_NHT Homo sapiens cdna clone IMAGE:1409245
3' similar to SW:GGE2_HUMAN Q13066 GAGE-2 PROTEIN. [1] ;, mRNA
sequence.
ACCESSION AA868226
VERSION AA868226.1 GI:2963671
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 426)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo
, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Possible reversed clone: similarity on wrong strand
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 363.
Location/Qualifiers
1. 426
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1409245"
/clone_lib="Soares_testis_NHT"
/sex="male"
/lab_host="DH10B"
/notes="Vector: pMT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from Clontech Laboratories
, Inc., and primed with a Not I - oligo(dT) primer [5',
TGTTACCAATCTGAAGTGGAGCGGCCCAATTTTGTTTT 3'] .
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pMT73 vector. Library
went through one round of normalization to Cot5, and was
constructed by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT 79 a 119 c 95 g 131 t
ORIGIN
source
Query Match 76.7%; Score 412.8; DB 9; Length 426;
Best Local Similarity 97.9%; Pred. No. 9.4e-86;
Matches 417; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 102 TCGACCTATCGGCTAGACCAAGACGCTACGTAGAGCCTCTCGAATGATTGGCCTATG 161
|||||
Db 426 TCGACCTATCGGCTAGACCAAGACGCTACGTAGAGCCTCTCGAATGATTGGCCTATG 367
|||||
```

Qy 162 .CGGCCGAGCAGTTCAGTCATGAAGTGGAAACACAGCACACCTGAAGAGGGGAACACGCA 221
|||||
Db 366 CGGCCGAGCAGTTCAGTCATGAAGTGGAAACACAGCACACCTGAAGAGGGGAACACGCA 307
|||||
Qy 222 ACTCAACGTCAGGATCCTCAGTCGCTCAGGAGGGAGGATGAGGGAGCATCTGCAGGT 281
|||||
Db 306 ACTCAACGTCAGGATCCTCAGTCGCTCAGGAGGGAGGATGAGGGAGCATCTGCAGGT 247
|||||
Qy 282 CAAGGCCGAGCAGTTCAGTCATGAAGTGGAAACACAGCACACCTGAAGAGGGGAACACG 341
|||||
Db 246 CAAGGCCGAGCAGTTCAGTCATGAAGTGGAAACACAGCACACCTGAAGAGGGGAACACG 187
|||||
Qy 342 TGTGAAGTGGTCTCAGTGGCAGGAGATGACCGCCAAATCCAGAGGAGGTGAAACG 401
|||||
Db 186 TGTGAAGTGGTCTCAGTGGCAGGAGATGACCGCCAAATCCAGAGGAGGTGAAACG 127
|||||
Qy 402 CTTGAAGAGGTGAAAGCAATCACAGTGTAAAGAGCACACCTTGAATGATGCAGGC 461
|||||
Db 126 CTTGAAGAGGTGAAAGCAATCACAGTGTAAAGAGCACACCTTGAATGATGCAGGC 67
|||||
Qy 462 TGTCTCTATGTTGGAATTTGTTCAATTAATAATCTCCCAATAAAAGCTTTACAGCCTTCTG 521
|||||
Db 66 TGTCTCTATGTTGGAATTTGTTCAATTAATAATCTCCCAATAAAAGCTTTACAGCCTTCTT 7
|||||
Qy 522 CAAAGA 527
|||||
Db 6 CAAAAA 1
|||||
RESULT 10
AA760996/c
LOCUS
DEFINITION
n32h08_s1 NCI_CGAP_G04 Homo sapiens cDNA clone IMAGE:1257855 3'
similar to TR:Q13070 Q13070 GAGE-6 PROTEIN. ;, mRNA sequence.
ACCESSION
AA760996
VERSION
AA760996.1 GI:2809926
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 450)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
JOURNAL
COMMENT
Contact: cgapb-r@mail.nih.gov
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
Emmert-Buck, M.D., Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 331.
FEATURES
source
1..450
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1257855"
/clone_lib="NCI_CGAP_G04"
/tissue_type="pooled germ cell tumors"
/lab_host="DH10B"
/note="Vector: pRT3D-Pac (Pharmacia) with a modified
polylinker; 1st strand cDNA was prepared from 3 pooled
germ cell tumors, and was then primed with a Not I -
oligo(dT) primer. Double-stranded cDNA was ligated to Eco
RI adaptors (Pharmacia), digested with Not I and cloned
into the Not I and Eco RI sites of the modified pRT3
vector. Library is normalized. Library was constructed by

BASE COUNT 85 a 130 c 94 g 140 t 1 others
ORIGIN
Query Match 76.4%; Score 410.8; DB 9; Length 450;
Best Local Similarity 97.3%; Pred. No. 2.7e-85;
Matches 439; Conservative 0; Mismatches 8; Indels 4; Gaps 2;
Qy 84 ATCAGTTGGCGAGGAAGATCGACC---TATCGGCCTAGACCAAGACGCTACGTAGAGCCT 140
|||||
Db 450 ATGAGTTGGCGAGGAAGATCGACCCTATTATTGGCCTAGACCAAGGCGCTATGTACAGCCT 391
|||||
Qy 141 CCTGAAATGATTTGGCCCTATCGGCCCGCAGCAGTTCAGTGATGAACTGGAACACGCAACA 200
|||||
Db 390 CCTGAAGTGAATGTGCTTATCGGCCCGCAGCAGTTCAGTGATGAACTGGAACACGCAACA 331
|||||
Qy 201 CCTGAAAGAGGGGAACACAGCAACTCAAGCTCAGGATCCTCGAGTCTCTCAGGAGGGAGAG 260
|||||
Db 330 CCTGAAAGAGGGGAACACAGCAACTCAAGCTCAGGATCCTCGAGTCTCTCAGGAGGGAGAG 271
|||||
Qy 261 GATGAGGAGCATCTGCAGGTCAAGGGCCGAAAGCCTGAAGCTCATAGCCAGGAACAGGGT 320
|||||
Db 270 GATGAGGAGCATCTGCAGGTCAAGGGCCGAAAGCCTGAAGCTCATAGCCAGGAACAGGGT 211
|||||
Qy 321 CACCCACAGACTGGGTGTGAGTGTGAAGATGGTCTGTATGGGCGAGAGATGGACCCGCCA 380
|||||
Db 210 CACCCACAGACTGGGTGTGAGTGTGAAGATGGTCTGTATGGGCGAGAGATGGACCCGCCA 151
|||||
Qy 381 AATCCAGAGGAGGTGAAAGCGCTGAAGAGGTGAAAGCAATCACAGTCTTAAAGAG 440
|||||
Db 150 AATCCAGAGGAGGTGAAAGCGCTGAAGAGGTGAAAGCAATCACAGTCTTAAAGAG 91
|||||
Qy 441 ACAGTTGAAATGATGACAGGCTGCTCTATGTTGGAAATTTGTCATTAATTTCTCCCA 500
|||||
Db 90 ACAGTTGAAATGATGACAGGCTGCTCTATGTTGGAAATTTGTCATTAATTTCTCCCA 31
|||||
Qy 501 ATAAGCTTTACAGCTTCTGCAAGAGAAAAA 531
|||||
Db 30 ATAAG-TTTACAGCTTCTGCAAGAGAAAAA 1
|||||
RESULT 11
AI187350/c
LOCUS
DEFINITION
qf29a05.x1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:1751408
3' similar to SW:GGE4_HUMAN Q13068 GAGE-4 PROTEIN. [1] ;, mRNA
sequence.
ACCESSION
AI187350
VERSION
AI187350.1 GI:3737988
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 455)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
JOURNAL
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapb-r@mail.nih.gov
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo
Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert length: 504 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 391.
FEATURES
Location/Qualifiers

```
source
1. .455
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1751408"
/clone_lib="Soares_testis_NHT"
/sex="male"
/lab_host="DH10B"
/notes="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from Clontech Laboratories
Inc., and primed with a Not I - oligo(dT) primer [5'
TGTTACCAATCTGAAGTGGAGCGCGCCCAATTTTITTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization to Cot5, and was
constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT      86 a 130 c 95 g 144 t
ORIGIN

Query Match      74.6%; Score 401.2; DB 9; Length 455;
Best Local Similarity 95.4%; Pred. No. 4.6e-83;
Matches 435; Conservative 0; Mismatches 18; Indels 3; Gaps 2;

QY 84 ATGAGTTGGCGAGGAAGATCGACCTAT--CGGCGCTAGACCAAGAGCGTACGTAGAGCCCTC 141
|||||
DB 455 ATGAGTTGGCGAGGAAGATGACCTTATATTGGCTAGACAGGCGCTATGT-CAGCCCTC 397
|||||

QY 142 CTGAATGATTTGGGCTATGCGGCCGCGAGCAGTTCACTGATGAAGTGAACCAACAC 201
|||||
DB 396 CTGAATGATTTGGGCTATGCGGCCGCGAGCAGTTCACTGATGAAGTGAACCAACAC 337
|||||

QY 202 CTGAAGAAGGGAACAGCAACTCAAGTCAGGATCTCGAGCTGAGCTGAGGAGGAGG 261
|||||
DB 336 CTGAAGAAGGGAACAGCAACTCAAGTCAGGATCTCGAGCTGAGGAGGAGGAGG 277
|||||

QY 262 ATGAGGAGCATCTGAGGTCAGGCGCGAGCAGTTCACTGATGAAGTGAACCAACAGGCTC 321
|||||
DB 276 ATGAGGAGCATCTGAGGTCAGGCGCGAGCAGTTCACTGATGAAGTGAACCAACAGGCTC 217
|||||

QY 322 ACCACAGACTGGGTGTGAGTGTGAAGATGCTCTGTATGGCAGGAGATGACCGCCCAA 381
|||||
DB 216 ACCACAGACTGGGTGTGAGTGTGAAGATGCTCTGTATGGCAGGAGATGACCGCCCAA 157
|||||

QY 382 ATCCAGAGAGGTGAAAACCGCTGAAAGAGGTGAAAAGCAATCACAGTGTTTAAAGAAGA 441
|||||
DB 156 ATCCAGAGAGGTGAAAACCGCTGAAAGAGGTGAAAAGCAATCCACAGTGTTTAAAGAAGA 97
|||||

QY 442 CACGTTGAATGATGACAGCTGCTCCTATCTTGTGAAATTTGTTCAATAAATTTCTCCCAA 501
|||||
DB 96 CACGTTGAATGATGACAGCTGCTCCTATCTTGTGAAATTTGTTCAATAAATTTCTCCCAA 37
|||||

QY 502 TAAAGCTTTACGCCCTTCTGCAAGAAGAAAAA 537
|||||
DB 36 TAAAGCTTTCCAGCCTTCTGCAAAAAA 1
|||||

RESULT 12
AW102587/c AW102587 412 bp mRNA linear EST 19-OCT-1999
LOCUS      xd67g07.x1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone
DEFINITION IMAGE:2602716 3' similar to SW:GGE2_HUMAN Q13066 GAGE-2 PROTEIN. [1
            ]; mRNA sequence.
ACCESSION  AW102587
VERSION     AW102587.1 GI:6073200
KEYWORDS    EST..
SOURCE      Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 412)
AUTHORS     NGI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
```

```
TITLE      National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
JOURNAL     Tumor Gene Index
COMMENT     Unpublished (1997)
            Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-re@mail.nih.gov
            This clone is available royalty-free through LINL; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -40UP from Gibco
            High quality sequence stop: 285.

FEATURES
            Location/Qualifiers
            1..412
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="IMAGE:2602716"
            /clone_lib="Soares_NFL_T_GBC_S1"
            /lab_host="DH10B"
            /notes="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
            a modified polylinker; Site 1: Not I; Site 2: Eco RI;
            Equal amounts of plasmid DNA from three normalized
            libraries (fetal lung NBHL19W, testis NHT, and B-cell
            NCI_CGAP_GCB1) were mixed, and ss circles were made in
            vitro. Following HAP purification, this DNA was used as
            tracer in a subtractive hybridization reaction. The driver
            was PCR-amplified cDNAs from pools of 5,000 clones made
            from the same 3 libraries. The pools consisted of
            I.M.A.G.E. clones 297480-302087, 682632-687239,
            726408-728711, and 729096-731399. Subtraction by Bento
            Soares and M. Fatima Bonaldo."
            BASE COUNT      78 a 114 c 84 g 133 t
            ORIGIN

Query Match      67.9%; Score 365.4; DB 9; Length 412;
Best Local Similarity 97.8%; Pred. No. 9.6e-75;
Matches 391; Conservative 0; Mismatches 7; Indels 2; Gaps 2;

QY 138 CCTCCTGAATGATTTGGGCTTATGCGGCCCGAGCAGTTCAGTGAAGTGAACCA-GC 196
|||||
DB 400 CCTCTTTGAATGATTTGGGCTTATGCGGCCCGAGCAGTTCAGTGAAGTGAACCAANGC 341
|||||

QY 197 AACACCTTGAAGAGGGGAACACAG-CAACTCAACGTCAGGATCCTGCGAGTCTGCAGGAGG 255
|||||
DB 340 AACACTTGAAGAGGGGAACACAGCACTCANAGTCAGGATCCTTCAGCTGCTCAGNAGG 281
|||||

QY 256 GAGAGATGAGGGAGCATCTCCAGGTCAAGGCCCGAGCCTGAAGCTCATAGCCAGGAAC 315
|||||
DB 280 GAGAGATGAGGGAGCATCTCCAGGTCAAGGCCCGAGCCTGAAGCTCATAGCCAGGAAC 221
|||||

QY 316 AGGTCACCCACACACTGGGTGTGAGTGTGAAGATGGTCTGATGGCAGGAGATGGACC 375
|||||
DB 220 AGGTCACCCACACACTGGGTGTGAGTGTGAAGATGGTCTGATGGCAGGAGATGGACC 161
|||||

QY 376 CGCCAAATCCAGAGGAGGTGAAAACCGCTGAAGAGGTGAAAAGCAATCACAGTGTAAA 435
|||||
DB 160 CGCCAAATCCAGAGGAGGTGAAAACCGCTGAAGAGGTGAAAAGCAATCACAGTGTAAA 101
|||||

QY 436 AGAAGACACCTTGAAATGATGCAGCTGCTCCTATGTTGGAATTTGTTCAATAAATTC 495
|||||
DB 100 AGAAGACACCTTGAAATGATGCAGCTGCTCCTATGTTGGAATTTGTTCAATAAATTC 41
|||||

QY 496 TCCCAATAAGCTTTACAGCCTTCTGCAAGAAGAAAAA 535
|||||
DB 40 TCCCAATAAGCTTTACAGCCTTCTGCAAAAAA 1
|||||

RESULT 13
BC005363
LOCUS      BC005363
DEFINITION Homo sapiens, clone IMAGE:3686872, mRNA.
ACCESSION  BC005363
VERSION     BC005363.1 GI:14710126
KEYWORDS    HTC.
SOURCE      human.
```

ORGANISM: Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 445)
 Strausberg, R.
 Direct Submission
 Submitted (27-MAR-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgapbs-re@mail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: CLONTECH Laboratories, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome
 Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www-shgc.stanford.edu>
 Contact: (Dickson, Mark) mcd@paxil.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
 R. M.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LNLN at: <http://image.llnl.gov>
 Series: IRAL Plate: 16 Row: e Column: 15
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 4503882
 This clone has the following problem: frame shifted.

FEATURES
 Location/Qualifiers
 1..445
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:3686872"
 /tissue_type="Bone marrow, chronic myelogenous leukemia"
 /clone_lib="NIH_MGC_54"
 /lab_host="DH10B"
 /note="Vector: pDNR-LIB"

BASE COUNT 127 a 102 c 132 g 84 t

ORIGIN

Query Match 67.0%; Score 360.4; DB 11; Length 445;
 Best Local Similarity 95.3%; Pred. No. 1.4e-73;
 Matches 383; Conservative 0; Mismatches 16; Indels 3; Gaps 1;

Qy 1 ACGCCAGGAGCTGTGAGGCAGTGTCTGTGGTTCCTGCCGTCGGACATCTTTTCCTCT 60
 Db 35 ACGCCAGGAGCTGTGAGGCAGTGTCTGTGGTTCCTGCCGTCGGACATCTTTTCCTCT 94
 Qy 61 ACTGAGATTCATCTGTGTAATATGATGTCGCGAGGAAGATCGACC---TATCGGCCTA 117
 Db 95 ACTGAGATTCATCTGTGTAATATGATGTCGCGAGGAAGATCGACCTTTATTGGCCTA 154
 Qy 118 GACCAAGACGCTACGTAGAGCCCTCCTGAAATGATTGGGCCTATGGGGCCCGAGCAGTTCA 177
 Db 155 GTCCAAGACGCTATGTACAGCCTCCTGAAATGATTGGGCCTATGCAGGCCGAGCAGTTCA 214
 Qy 178 GTGATGAAGTGGAACACGACCAACCTGAGAGGGGAACCACTCAACGTGAGATC 237
 Db 215 GTGATGAAGTGGAACACGACCAACCTGAGAGGGGAACCACTCAACGTGAGATC 274
 Qy 238 CTGCAGCTGCTCAGGAGGAGAGGATGAGGGAGCATCTGCAGGTCAAGGGCCGAAGCCTG 297
 Db 275 CTGCAGCTGCTCAGGAGGAGAGGATGAGGGAGCATCTGCAGGTCAAGGGCCGAAGCCTG 334
 Qy 298 AAGCTCTATAGCCAGGAACAGGGTCCACCCACAGACTGGGTGTGAGTGTGAAGATGGTCTG 357
 Db 335 AAGCTGTATAGCCAGGAACAGGGTCCACCCACAGACTGGGTGTGAGTGTGAAGATGGTCTG 394
 Qy 358 ATGGCAGGAGATGGACCCGCCAATCCAGAGGAGGTGAAAA 399
 Db 395 ATGGGCAGGAGATGGACCCGCCAATCCAGAGGAGGTGAAAA 436

RESULT 14	AA448542	419 bp	mrna	linear	EST 04-JUN-1997
LOCUS	AA448542				
DEFINITION	zw91ell.r1 Soares.testis.NHT Homo sapiens cDNA clone IMAGE:782636				
	5', similar to TR:G914905 G914905 GAGE-4 PROTEIN. [1] ;, mRNA				
	sequence.				
ACCESSION	AA448542				
VERSION	AA448542.1	GI:2162212			
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.				
	1 (bases 1 to 419)				
	Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,				
	Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B.				
	, Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wyllie				
	,T., Waterston,K., and Wilson,R.				
TITLE	WASHU-Merck EST Project 1997				
JOURNAL	Unpublished (1997)				
COMMENT	Contact: Wilson RK				
	Washington University School of Medicine				
	4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108				
	Tel: 314 286 1800				
	Fax: 314 286 1810				
	Email: est@watson.wustl.edu				
	This clone is available royalty-free through LLNL ; contact the				
	IMAGE Consortium (info@image.llnl.gov) for further information.				
	Seq primer: -28ml3 rev2 ET from Amersham				
	High quality sequence stop: 353.				
FEATURES	Location/Qualifiers				
source	1. .419				
	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
	/clone="IMAGE:782636"				
	/clone_lib="Soares_testis_NHT"				
	/sex="male"				
	/lab_host="DH10B"				
	/note="Vector: p7T3D-Pac (Pharmacia) with a modified				
	polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA				
	was prepared from mRNA obtained from Clontech Laboratories				
	, Inc., and primed with a Not I - oligo(dT) primer [5',				
	TGTTGACCAATCTGAAGTGGAGCGCGCCCAATTTTTTTTTTTT 3'].				
	Double-stranded cDNA was ligated to Eco RI adaptors				
	(Pharmacia), digested with Not I and cloned into the Not I				
	and Eco RI sites of the modified p7T3 vector. Library				
	went through one round of normalization to Cot5, and was				
	constructed by Bento Soares and M. Fatima Bonaldo. "				
BASE COUNT	108 a	98 c	134 g	79 t	
ORIGIN					
Query Match	66.8%; Score 359.4; DB 9; Length 419;				
Best Local Similarity	96.3%; Pred. No. 2.4e-73;				
Matches 390; Conservative	0; Mismatches 11; Indels 4; Gaps 2;				
Qy	1	ACGCCAGGAGCTGTGAGGCAGTCTGCTGTGGTTCCTGCCGTCGCCGACACTCTTTTTCCTCT	60		
Db	16	ACGCCAGGAGCTGTGAGGCAGTCTGCTGTGGTTCCTGCCGTCGCCGACACTCTTTTTCCTCT	75		
Qy	61	ACTGAGATTCATCTGCTGAAATATGAGTTGGCGAGGAAGATCGACC---	117		
Db	76	ACTGAGATTCATCTGCTGAAATATGAGTTGGCGAGGAAGATCGACCTATATTGGCCTA	135		
Qy	118	GACCAAGAGCCTACGTAGAGCCTCCTGAAATGATTGGGCTTATGCGGCCGACAGTTCA	177		
Db	136	GACCAAGAGCCTATGTACAGCCTCCTGAAGTATTGGGCTTATGCGGCCGACAGTTCA	195		
Qy	178	GTGATGAAGTGGAAACAGCAACACTGAAGAGGGGAACAGCAACTCAACGTAGGATC	237		
Db	196	GTGATGAAGTGGAAACAGCAACACTGAAGAGGGGAACAGCAACTCAACGTAGGATC	255		

```
QY 238 CTGACGTGCTCAGGAGGAGGATGAGGAGGAGCATCTCCAGGTCAAGGCCGGAAGCCTG 297
|||||
Db 256 CTGCAGCTGCTCATCGCGGAGGAGGATGAGGAGGAGCATCTGCAGTCAAGGCCGGAAGCCTG 315
|||||
QY 298 AAGCTCATAGCCAGAACAGGGTCAACCACAGACTGGGTGTGAGTGTGAAGATGGTCCCTG 357
|||||
Db 316 AAGCTCATAGCCAGAACAGGGTCA-CCACAGACTGGGTGTGAGTGTGAAGATGGTCCCTG 374
|||||
QY 358 ATGGGCAGGAGATGGACCCGCCAAATCCAGAGGAGGTGAAAACGC 402
|||||
Db 375 ATGGGCAGGAGATGGACCCGCCAAATCCAGAGGAGGTGAAAACGC 419
|||||

RESULT 15
AA913206/c
LOCUS
DEFINITION
om53f03.s1 NCI-CGAP_GC4 Homo sapiens cDNA clone IMAGE:1550909 3'
similar to SW:GGE3_HUMAN Q13067 GAGE-3 PROTEIN. [1] ; , mRNA
sequence.
ACCESSION
AA913206
VERSION
AA913206.1 GI:3052598
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 384)
AUTHORS
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL
Unpublished (1997)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 487 Std Error: 0.00
Seq Primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 81.
FEATURES
source
1..384
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1550909"
/tissue_type="NCI-CGAP_GC4"
/lab_host="DH10B"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; 1st strand cDNA was prepared from 3 pooled
germ cell tumors, and was then primed with a Not I -
oligo(dT) primer. Double-stranded cDNA was ligated to Eco
RI adaptors (Pharmacia), digested with Not I and cloned
into the Not I and Eco RI sites of the modified pT7T3
vector. Library is normalized. Library was constructed by
Bento Soares and M. Fatima Bonaldo."
BASE COUNT 63 a 119 c 83 g 119 t
ORIGIN

Query Match 61.3%; Score 329.8; DB 9; Length 384;
Best Local Similarity 94.0%; Pred. No. 1.8e-66;
Matches 343; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 100 GATCGACCTATCGGCTAGACCAAGAGCGTACGTAGAGCCCTCCCTGAAATGATTGGCCCTA 159
|||||
Db 365 GACCTATTATTGGCCCTAGACCAAGGGCGCTATGTACAGCCCTCTCGAAATGAAATGGCCTA 306
|||||
```

```
QY 160 TCGGCCCCGAGCAGATTTCAGTGTGAAGTGGAAACCAACACACCTGTAAGAGGGGAACACG 219
|||||
Db 305 TCGGCCCCGGAACAGTTTCAGGGATGAAGTGGAAACCAACACACCTGTAAGAGGGGAACACG 246
|||||
QY 220 CAACTCAACGTTCAGGATCCCTGCAGCTGCTCAGGAGGGAGAGGATGAGGGAGCATCTCCAG 279
|||||
Db 245 CAACTCAACGTTCAGAAATCCCTGCAGCTGCTCAGGAGGGAGAGGATGAGGGAGCATCTCCAG 186
|||||
QY 280 GTCAAGGGCCGGAAGCCCTGAAGCTCATAGCCAGGAAACAGGGTCACCCACAGACTGGGTGTG 339
|||||
Db 185 GTCAAGTGGCCGAAGCCCTGAAGCTCATAGCCAGGAAACAGGGTCACCCACAGACTGGGTGTG 126
|||||
QY 340 AGTGTGAAGATGGTCTCTGATGGGCAGGAGATGGACCCGCCAAATCCAGAGGAGGTGAAAA 399
|||||
Db 125 AGTGTGAAGATGGTCTCTGATGGGCAGGAGATGGACCCGCCAAATCCAGAGGAGGTGAAAA 66
|||||
QY 400 CGCCTGAAGAAAGTGAAGAAACCAATCACAGTCTTAAAGAGAGACACGTTTGAATGATCCAG 459
|||||
Db 65 CGCCTGAAGAAAGTGAAGAAACCAATCACAGTCTTAAAGAGAGACACGTTTGAATGATCCAG 6
|||||
QY 460 GCTGC 464
|||||
Db 5 GCTGC 1
|||||
```

Search completed: July 1, 2002, 07:07:18
Job time: 1784 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 1, 2002, 08:39:24 ; Search time 177.44 seconds
(without alignments)
5205.703 Million cell updates/sec

Title: US-09-782-745-14
Perfect score: 538
Sequence: 1 ACGCAGGAGCGTGTGAGGC.....CTGCAAGAGAAAAA 538

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues
Total number of hits satisfying chosen parameters: 2825752

Minimum DB seq length: 0
Maximum DB seq length: 538

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_032802.*

- 1: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
- 2: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
- 3: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
- 4: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
- 5: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
- 6: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
- 7: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
- 8: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
- 9: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
- 10: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
- 11: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
- 12: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
- 13: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
- 14: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
- 15: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
- 16: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
- 17: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
- 18: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
- 19: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
- 20: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
- 21: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
- 22: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
- 23: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
- 24: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Score	Match Length	DB ID	Description
1	538	100.0	538 20 AAX90519	GAGE-2 tumour reje
2	504.8	93.8	532 20 AAX90522	GAGE-5 tumour reje
3	496.8	92.3	532 19 AAV18720	CDNA encoding GAGE
4	493	91.6	535 19 AAV18717	CDNA encoding GAGE
5	357.8	66.5	430 21 AAC02129	Human secreted pro
6	301.4	56.0	530 22 AAS60104	Human cancer agent
C	299.2	55.6	22 AAS60496	Human cancer agent
8	194	36.1	532 22 AAT60530	Human polynucleoti
9	172.4	32.0	503 23 AAS69484	DNA encoding novel

10	130.4	24.2	475	22	AAD14981	Human NOV2 DNA. H
11	108	20.1	399	22	AAF68151	Human lung tumour
12	101.6	18.9	515	22	AAS63900	Human prostate cDN
13	101.6	18.9	515	22	AAH93807	Human prostate-spe
14	101.6	18.9	515	22	AAH85121	Human prostate-spe
15	101.6	18.9	515	22	AAH02872	prostate tumour an
16	100.8	18.7	320	22	AAS24637	Human ovarian PCR-
17	100.8	18.7	320	22	AAH83260	Human ovarian tumo
C	93.4	17.4	472	21	AAZ95012	CSG Prol18 EST use
18	93.4	17.4	476	21	AAZ95012	PAGE-4 coding sequ
19	93.4	17.4	478	21	AAZ95013	Cancer specific ge
20	92.8	17.2	457	22	AAF68851	Human lung tumour
21	92.8	17.2	420	21	AAA06691	Human immunogenic
22	92.4	16.1	348	22	AAD14997	PCR2.1-83350421-S7
C	84.8	15.8	479	22	AAF68152	Human lung tumour
23	82.8	15.4	458	22	AAH14980	Human NOV1 DNA. H
24	82.8	15.4	458	22	AAH14980	Human secreted pro
25	79.2	14.7	372	20	AAX41280	Human lung tumour
26	77.4	14.4	461	22	AAF68852	EST clone BN365.
27	75.4	14.0	328	20	AAV87256	Human cDNA clone F
28	68.4	12.7	300	20	AAX55997	Human secreted pro
C	53	9.9	157	21	AAC15593	Human secreted exp
30	53	9.9	201	21	AAA42613	Human DNA containi
C	32	9.5	51	22	AAH79806	Human SNP oligonuc
33	49.4	9.2	51	22	AAL33515	Human DNA containi
34	49.4	9.2	51	22	AAH79998	Human SNP oligonuc
35	46.2	8.6	51	22	AAL32796	Human DNA containi
36	44.6	8.3	51	22	AAH79997	Human DNA containi
37	44.2	8.2	51	22	AAL27748	Human SNP oligonuc
38	40	7.4	372	23	AAS75452	DNA encoding novel
39	38.2	7.1	51	22	AAL27747	Human SNP oligonuc
40	38	7.1	51	22	AAL32276	Human SNP oligonuc
41	37.8	7.0	498	21	AAC45680	Arabidopsis thalia
42	37.8	7.0	500	21	AAC39188	Arabidopsis thalia
43	37.4	7.0	50	22	AAL32275	Human SNP oligonuc
44	37.2	6.9	237	21	AAC66318	ETAV tat gene SEQ
45	36.6	6.8	501	23	AAS66773	DNA encoding novel

ALIGNMENTS

RESULT 1
AAX90519
ID AAX90519 standard; cDNA; 538 BP.
XX
AC AAX90519;
XX
DT 30-SEP-1999 (first entry)
DE GAGE-2 tumour rejection antigen clone nucleotide sequence.
XX
KW Human leukocyte antigen; HLA-A29; tumour rejection antigen;
KW detection; therapy; pathological condition; cancer; CTL;
KW cytolytic T lymphocyte; GAGE; ss.
XX
OS Homo sapiens.
XX
PN WO9937665-AL.
XX
PD 29-JUL-1999.
XX
PF 12-JAN-1999; 99WO-US00775.
XX
PR 23-JAN-1998; 98US-0012818.
XX
(LUDW-) LUDWIG INST CANCER RES.
XX
PI Boon-Falleur T, Debacker O, Van Den Eynde B, Van Der Bruggen P;
XX WPI; 1999-469111/39.
XX
PT New isolated peptides which bind to HLA-A29 molecules, which are
PT tumour rejection antigens used for detection and therapy of


```
XX The present sequence encodes a GAGE-2 tumour rejection antigen
CC precursor (TRAP). The protein is expressed in a number of tumours. In
CC contrast the only normal tissue which expresses GAGE TRAP protein is
CC testis. Several GAGE TRAPS have been identified (see AAV18/17-21). The
CC major difference between these proteins and GAGE-1 (AAV05540) is the
CC absence of a stretch of 143 bases located at position 379 to 521 of the
CC GAGE-1 TRAP sequence. The rest of the sequences show mismatches at
CC various position, with the exception of GAGE-3 whose 5' end is totally
CC different from the other GAGE cDNAs for the first 112 bases. This
CC region of GAGE-3 cDNA contains a long repeat and a hairpin structure.
CC The antigens can be used to diagnose melanomas, characterised by
CC expression of a TRAP or presentation of a tumour rejection antigen.
CC Antigens shed into blood or urine can be observed and then used to
CC confirm a diagnosis of melanoma using cytolytic T cell clone
CC proliferation methodologies. Other uses for the processed peptides,
CC include HLA-typing assays for, e.g. skin graft or organ transplants.
XX
SQ Sequence 535 BP; 158 A; 112 C; 156 G; 109 T; 0 other;

Query Match          91.6%; Score 493; DB 19; Length 535;
Best Local Similarity 97.2%; Pred. No. 3.5e-133;
Matches 524; Conservative 0; Mismatches 10; Indels 5; Gaps 2;

QY 1 ACGCCAGGGAGCTGTGAGGCAGTCTGTGTGGTTCTCTGCCGCCGACTCTTTTTCCTC- 59
DB 1 acgcccaggagcgtgtgaggcagtgctgtg----ttcttgcgtcggactcttttctct 56
QY 60 TACTGAGATCACTGTGTGAAATGATGAGTGGCGAGGAAGATCGACCTATCGCCCTAGA 119
DB 57 tactgagattcatctgtgaaatgattggtggcgaggagatcgacctatcgccctaga 116
QY 120 CCAAGACGCTACGTAGAGCCCTCCGAAATGATTGGGCCCTATGGCCCGCAGCAGTTCA 179
DB 117 ccaagacgtactagagcctcctgaaatgattggccctatgcggcccgagcagttcagt 176
QY 180 GATGAAGTGAACCAACCTGTGAAGAGGGGGAACACAGCACTCAACGTTCAGGATCCT 239
DB 177 gatgaagtgaaccagagcctcctgaagaagggggaaccagcaactcaacgtcaggatct 236
QY 240 GCAGCTGCTCAGAGGGAGAGGATGAGGAGCATCTGCAGGTCAAGGCCGAGCCTGAA 299
DB 237 gcagctgctcagagggagagagatgagggagcatctgcaggtcaaggccgaagcctgaa 296
QY 300 GCTCATAGCCAGGAACAGGGTCAACCCACAGACTGGGTGTGAGTGTGAAGATGGTCTGAT 359
DB 297 gctgagagccaggaacagggtccaccacagactgggtgtgagtgtgaagatggtcctgat 356
QY 360 GGCAGGAGATGGACCCGCCAAATCCAGAGAGGTGAAACGGCTGAAGAGGTGAAGAAG 419
DB 357 gggcaggagatggaaccgcgcaaatccagaggaggtgaaacgcgctgaagaggtgaaag 416
QY 420 CAATCACAGTGTAAAGAAGACACAGTTCGAATGATGCGAGGCTGCTCCTATGTTGGAAT 479
DB 417 caatcacagtgtaaaagaagacatgttgaaatgatgcagggcgtcctatgttcggaat 476
QY 480 TTGTTTCATTAATAATTCCTCCAAATAAGCTTTACAGCCTTCTGCAAGAAAAAATAAAA 538
DB 477 ttgttcattaaattcctcccaataaagctttacagccttctgcaagaaaaaataaaaa 535

RESULT 5
AAC02129
ID AAC02129 standard; cDNA; 430 BP.
XX
AC AAC02129;
XX
DT 06-OCT-2000 (first entry)
XX
DE Human secreted protein 5' EST, SEQ ID NO: 2127.
XX
KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
```

```
gene therapy; chromosome mapping; ss.
XX Homo sapiens.
XX PN EP1033401-A2.
XX PD 06-SEP-2000.
XX PF 21-FEB-2000; 2000EP-0200610.
XX PR 26-FEB-1999; 99US-0122487.
XX PA (GEST ) GENSET.
XX PI Dumas Milne Edwards J, Duclert A, Giordano J;
XX WP1: 2000-500381/45.
XX P-PSDB; AAG02123.
XX
PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX Claim 1; SEQ ID 2127; 71pp + CD-ROM; English.
XX
CC The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. An ORF has been identified within the
CC sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs
CC derived from 30 different tissues. EST sequences usually correspond
CC mainly to the 3' untranslated region (UTR) of the mRNA because they are
CC often obtained from oligo-dr primed cDNA libraries. Such ESTs are not
CC well suited for isolating cDNA sequences derived from the 5' ends of
CC mRNAs and even in those cases where longer cDNA sequences have been
CC obtained, the full 5' UTR is rarely included. 5' ESTs are derived from
CC mRNAs with intact 5' ends and can therefore be used to obtain full length
CC cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
CC gene therapy and chromosome mapping procedures. They are used to obtain
CC upstream regulatory sequences and to design expression and secretion
CC vectors.
XX
SQ Sequence 430 BP; 101 A; 102 C; 134 G; 85 T; 8 other;

Query Match          66.5%; Score 357.8; DB 21; Length 430;
Best Local Similarity 97.6%; Pred. No. 5.9e-94;
Matches 368; Conservative 5; Mismatches 2; Indels 2; Gaps 1;

QY 1 ACGCCAGGGAGCTGTGAGGCAGTCTGTGTGGTTCTCTGCCGCCGACTCTTTTTCCTCT 60
DB 54 acgccaggagcgtgtgaggcagtgctgtgtgttctctgcgctcgagactcttttctctct 113
QY 61 ACTGAGATTTCATCTGTGTGAAATATGAGTTGGCGAGGAAGATCGACCTAT--CGGCCCTAG 118
DB 114 actgagattcatctgtgaaatatgattggtgcgaggaagatcgacctatnagggcctag 173
QY 119 ACCAAGACGCTACGTAGAGCCTCTGTAATGATTGGGCTATGCGGCCCGCAGAGTTTCAG 178
DB 174 accaagrcgtatgtacagcctcctcctgaaatgattggcctatggtgcccgcgagagttcag 233
QY 179 TGATGAAGTGGGAACCAACAGCACACTGAAGAGGGGAACCCAGCAACTCAACGTCAGGATCC 238
DB 234 tgatgaagtggaaaccagcaacctgaagaagggggaaccagcaactcaaatgtcaggatcc 293
QY 239 TGCAGCTGCTCAGAGGGAGAGGATGAGGAGCATCTGCAGGTCAAGGGCCGAGCCTGA 298
DB 294 tgcagctgctcagragggagagagatgaggagcatctgcaggtcaaggccgaagcctga 353
QY 299 AGCTCATAGCCAGGAACAGGGTCACCCACAGACTGGGTGTGAGTGTGAAGATGGTCTCTGA 358
DB 354 agctcatagccaggaacagggtcaccaccagactgggtgtgagtgtgaagatggtcctga 413
QY 359 TGGGCGAGGAGATGGACC 375
DB 359 TGGGCGAGGAGATGGACC 375
```

Db	414	tgggcaggagatggamc	430	Db	389	ATCCTGNAGTCGNCAGNAGGGANAGGATNAGGNGCATNTGCAAGTCAAGGCCGCAAGC	330
RESULT	6			QY	295	CTGAAGCTCATAGCCAGGAACAGGGTCAACCCACAGACTGGGTGTGAGTGTGAAGATGGTC	354
AA560104/c				Db	329	CTGAAGNTGATAGCCAGGAACAGGGTCNCCCCANACTGGGNGTNAGTGTNAAATGGTC	270
XX	AA560104 standard; cDNA; 530 BP.			QY	355	CTGATGGCAGGAGATGGACCCGCCAAATCCAGAGGAGTGAAAACGGCTCTGAAGAAGTG	414
AC	AA560104;			Db	269	CTNATGGCAGNANATGGACCCNCCAAATCAAAGNAGGTGAAAACNCTCTGAANAAGTG	210
DT	29-JAN-2002 (first entry)			QY	415	AAAAGCAATCACAGTGTATAAAGAACACAGCTTGAAATGATGCAGGCTGCTCTATGTTG	474
XX	Human cancer agent-sensitive marker #105.			Db	209	AAAAGCAATCCAGTGTATAAANAAGCNCGTGAAATGATGCAGGCTGCTCTATGTTG	150
DE	Human cancer agent-sensitive marker #105.			QY	475	GAATTTTCTTCAATTAATTTCTCCCAATAAAGCTTTACAGCTTCTGCAAGAAAAAAA	534
XX	Human; cancer cell marker; TAXOL; cytostatic; tumour; carcinoma;			Db	149	GAATTTTCTTCAATTAATTTCTCCCAATAAAGCTTTACAGCTTNTNAAAAAANAANA	90
KW	squamous cell carcinoma; sarcoma; fibrosarcoma; leukaemia;			QY	535	AAAA 538	
KW	lymphocytic leukaemia; lymphoma; plasmocytoma; reticulum cell sarcoma;			Db	89	AAAA 86	
KW	Hodgkin's disease; glioma; ss.						
XX	Homo sapiens.						
OS	WO200179556-A2.						
PN	25-OCT-2001.						
XX	13-APR-2001; 2001WO-US12132.						
XX	14-APR-2000; 2000US-197538P.						
XX	(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.						
PI	Lillie J, Brown JL, Bolt A, Van Huffel C;						
XX	WPI; 2001-602933/68.						
XX	Novel nucleic acid, used as a marker to determine the effectiveness of						
PT	using TAXOL to treat cancer cell growth in individuals -						
XX	Claim 1; Page 120; 527pp; English.						
XX	The invention relates to 1046 novel nucleic acids which are used as						
CC	markers for determining the sensitivity of a cancer cell to the						
CC	anticancer agent TAXOL. Cancer cells can be treated with TAXOL when						
CC	they are shown to express one of the 242 sensitivity markers or the						
CC	cells are shown not to express one of the 804 resistance markers.						
CC	The methods can be used to determine the effectiveness of TAXOL						
CC	in the treatment of cancer cell growth in an individual. The markers						
CC	can be used as targets in developing anti-cancer agents such as						
CC	chemotherapeutic compounds. The markers can also be used as targets in						
CC	developing treatments for cancer, particularly those cancers which						
CC	display resistance to agents and exhibit expression of the markers. The						
CC	genomic sequences corresponding to the markers, in the identification of						
CC	cells or tissues which mis-express the protein. Cancers which may						
CC	be targeted include carcinoma (e.g. squamous cell carcinoma),						
CC	sarcoma (e.g. fibrosarcoma) leukaemia (e.g. lymphocytic leukaemia),						
CC	lymphoma, plasmocytoma, reticulum cell sarcoma, Hodgkin's disease and						
CC	tumours (e.g. glioma). The present sequence is one of the 1046						
CC	novel cancer cell markers.						
XX							
SQ	Sequence 530 BP; 89 A; 100 C; 106 G; 183 T; 52 other;						
	Query Match 56.0%; Score 301.4; DB 22; Length 530;						
	Best Local Similarity 87.9%; Pred. No. 1.6e-77;						
	Matches 320; Conservative 0; Mismatches 42; Indels 2; Gaps 1;						
QY	177 AGTGATGAAGTGAACACAGC--AACACCTGAAGAAGGGGACACCACTCAACCTCAGG	234					
Db	449 AGTATNAAGTGGAGCCACCCACNCCCTGAANAAGGNAACCACTCAACCTNAGG	390					
QY	235 ATCTGTCAGCTCTCAGGAGGAGATGAGGAGCATCTGCAGGTCAAGGCCGAGC	294					

CC cells or tissues which mis-express the protein. Cancers which may
CC be targeted include carcinoma (e.g. squamous cell carcinoma),
CC sarcoma (e.g. fibrosarcoma) leukaemia (e.g. lymphocytic leukaemia),
CC lymphoma, plasmocytoma, reticulum cell sarcoma, Hodgkin's disease and
CC tumours (e.g. glioma). The present sequence is one of the 1046
CC novel cancer cell markers.

XX
SQ Sequence 365 BP; 63 A; 86 C; 66 G; 150 T; 0 other;

Query Match 55.6%; Score 299.2; DB 22; Length 365;
Best Local Similarity 95.9%; Pred. No. 5.7e-77;
Matches 307; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 219 GCAACTCAACGTCAGGATCCTGCAGCTGCTCAGGAGGGAGAGATGAGGAGCATCTGCA 278
DB |||||
QY 365 GCAACTCAACGTCAGGATCCTGCAGCTGCTCAGGAGGGAAAGGATAAGGAGCATCTGCA 306
DB |||||
QY 279 GGTCAAGGCCGGAAGCTGAAGCTCATAGCCAGGAACAGGGFTACCCACAGACTGGGTGT 338
DB |||||
QY 305 GGTCAAGGCCGGAAGCTGAAGCTGATAGCCAGGAACAGGGFTACCCACAGACTGGGTGT 246
QY 339 GAGTGTGAAGATGCTCTGATGGCAGGAGATGGACCCGCCAAATCCAGAGGAGTGA 398
DB |||||
QY 245 GAGTGTGAAGATGCTCTGATGGCAGGAAATGGACCCGCCAAATCCAAAGAGGTGA 186
QY 399 ACGCTGGAAGAGGTGAAGAGCAATCACAGTGTAAAGAAAGACACAGCTTTGAAATGATGCA 458
DB |||||
QY 185 ACGCTGGAAGAGGTGAAGAGCAATCTCAGTGTAAAGAAAGGACGCTTGAAATGATGCA 126
QY 459 GCGTGTCTCTATGTTGAAATTTGTTTCAATTAATTTCCCAATAAAGCTTTACAGCCTT 518
DB |||||
QY 125 GCGTGTCTCTATGTTGAAATTTGTTTCAATTAATTTCCCAATAAAGCTTTACAGCCTT 66

QY 519 CTGCAAGAAAAA 538
DB 65 CTAAAAA 46

RESULT 8
AAI60530
ID AAI60530 standard; cDNA; 532 BP.
XX AC AAI60530;
XX DT 22-OCT-2001 (first entry)
XX DE Human polynucleotide SEQ ID NO 4519.
XX KW Human; neurotropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; ss.
XX OS Homo sapiens.
XX PN WO200153312-A1.
XX PD 26-JUL-2001.
XX PF 26-DEC-2000; 2000WO-US34263.
XX PR 21-JAN-2000; 2000US-0488725.
XX PR 25-APR-2000; 2000US-0552317.
XX PR 09-JUL-2000; 2000US-0598042.
XX PR 19-JUL-2000; 2000US-0620312.
XX PR 03-AUG-2000; 2000US-0653450.
XX PR 14-SEP-2000; 2000US-0662191.
XX PR 19-OCT-2000; 2000US-0693036.
XX PR 29-NOV-2000; 2000US-0727344.

PA (HYSE-) HYSEQ INC.
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX WPI: 2001-442253/47.
DR P-PSDB; AAM41374.
XX Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX Claim 1; SEQ ID NO 4519; 10078pp; English.
XX The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC the encoded polypeptides (AAM38642-AAM42213) with neurotropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
SQ Sequence 532 BP; 154 A; 108 C; 150 G; 120 T; 0 other;

Query Match 36.1%; Score 194; DB 22; Length 532;
Best Local Similarity 68.4%; Pred. No. 2.4e-46;
Matches 365; Conservative 0; Mismatches 14; Indels 24; Gaps 6;

QY 1 ACGCCAGGAGCTGTGAGGCAGTG-CTGTGTGGTTCCTGCCCTCGGACTCTTTTCCTC 59
DB |||||
QY 18 atgcccctggagtgagggtgagggtgcgttcctctgctgacttttctgctcc 77
DB |||||
QY 60 TACTGAGATTCATCTGTGAAATATAGTTGGGAGGAGATGACCTATCGCCCTAGA 119
DB |||||
QY 78 cactgagacgcagctgtgaaatatgattggcaggaagatcaacatagcctagg 137
QY 120 CCAAGACGCTACGTAGACCTCTGAAATGATTGGCCCTATCGCCGCCGAGCAGTTCAGT 179
DB |||||
QY 138 ccgaggagaagtgtaccacctctgagctgattggcctatgc-----tggagccggc 191
QY 180 GATGAAGTGGAAACACAGCAACCTGAAGAGGGGAACACGCAACTCAAGTCAGGATCCT 239
DB |||||
QY 192 gatgaggag-----cctcagcaagagagaccacccaactgaaagtgcggatcct 239
QY 240 GCAGTGTCTCAGGAG---GGAGAGGATGAGGAGGATCTGAGGTCAGGTCAGGCGGAGCCT 296
DB |||||
QY 240 gcacctggtcaggagagagaagatcagggtgcagctgagatcaadtgctgacctg 299
QY 297 GAAGCTCATAGCCAGGAACAGGCTCACCCACAGACTGGGTGTGAGTGTGAAGATGCTCCT 356
DB |||||
QY 300 gaagctgatctccaggagctgtctcagtaagaagctgggagtgatgagagatggctcct 359
QY 357 GATGGCAGGAGATGACCCGCCAAATCCAGAGAGGTGAAGAACCCCTGAA--GAAGGTGA 415
DB |||||
QY 360 gatgtccaggggaaagattctgacaaagtcagagcaatttaaaatccagaagggagtga 419
QY 416 AAGCAATCACAGTGTAAAGAGACACGTTGAAATGATGACGAGCTGCTCTATGTTGG 475
DB |||||
QY 420 caggcaaccacaggtttaaatgaagcagctgaacacacacacaaactgt-ttttatcta 478
QY 476 AAATTTGTTCAATTAATTTCTCCCAATAAAGCTTTTACAGCCTTTTCGCAAGAAAA 529
DB |||||
QY 479 agataattgacttaaaataatcgaataaaacttttgagcttttccgaaaaaaa 532

RESULT 9
AAS69484
ID AAS69484 standard; cDNA; 503 BP.
XX
AC AAS69484;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #5288.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
P1
DR WPI; 2001-639362/73.
DR P-PSDB; ABG05297.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity
XX
PS Claim 1; SEQ ID No 5288; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 503 BP; 158 A; 98 C; 121 G; 105 T; 21 other;

Query Match 32.0%; Score 172.4; DB 23; Length 503;
Best Local Similarity 68.2%; Pred. No. 4.4e-40;
Matches 307; Conservative 0; Mismatches 121; Indels 22; Gaps 4;
QY 76 TGTGAATATGAGTTGGCGAGGAAGATCGACCTATCGGCTAGACCAAGACGCTACCTAG 135
Db 52 tgtgaatatgtattggcgaggagatcaacatagctaggcctagcgagagagtgtac 111

QY 136 AGCCTCTGAAATGATTGGGCTATGCGCCGAGCAGTTCTAGTGATGAAGTGAACCCAG 195
Db 112 cactctgaagctgattggcctatgc-----tgagcccggtgatgagag----- 158
QY 196 CAACACCTGAAGAAGGGGAACAGCACTCAACGTACAGGATCTCGAGCTGCTCAGGAGG 255
Db 159 -----cctcagcaagaggaaccaccactgaaagtggatcggatcctgacctggtcaggaga 213
QY 256 G---AGAGGATGAGGAGCATCTGCAGGTCAAGGGCCGAAGCTCAAGCTCATAGCCAGG 312
Db 214 gaaagaagatcaggggttcagctgaagactcaagtgcctgacctgggaagctgactccaag 273
QY 313 AACAGGGTCAACCCACAGACTGGGTGTGAGTGTGAAGATGCTCTGTATGGCAGGAGATGG 372
Db 274 agctgtctcagtcaaaagactgggggtgaatgtgaaatggtcctgatgaccagggggaaga 333
QY 373 ACCGCGCAAAATCCAGAGAGGTGAACGCCCTGAAGAAGCTGAAGCAATTCACAGTGT 432
Db 334 ttctgccaataatcagaacaatttaaatccagaagagggtgacaggaaccacacaggttt 393
QY 433 AAAGAAGACACGCTTGAATGATGACGGCTGCTCTCTATGTTGGAATTTTGTTCATTA 492
Db 394 aaatgaagacaagctgaaacaccccaaacactgt-ttttattaagatatattgacttaaaa 452
QY 493 TTCCTCCCAATAAAGCTTTACAGCCTTCTGCG 522
Db 453 atatcgaaataaacttttcagcagctttctcc 482

RESULT 10
AAD14981
ID AAD14981 standard; DNA; 475 BP.
XX
AC AAD14981;
XX
DT 01-NOV-2001 (first entry)
XX
DE Human NOV2 DNA.
XX
KW Human; NOVX; G-antigen; GAGE-like protein; interferon;
KW G-protein coupled receptor; GPCR; hepatocyte nuclear factor;
KW mast cell protease; gene therapy; proliferative disorder; cancer;
KW immune disorder; hepatic disorder; cirrhosis; viral infection;
KW hepatitis; neuroendocrine system-related disorder; neurological disorder;
KW Parkinson's disease; infertility; autoimmune disease; arthritis;
KW multiple sclerosis; allergy; wound healing; cytostatic; nootropic;
KW immunosuppressive; neuroprotective; vulnery; hepatotropic; ds.
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FH 5'UTR 1..24
FT /*tag= a
FT CDS 25..360
FT /*tag= b
FT /*product= "Human NOV2 protein"
FT 3'UTR 361..475
FT /*tag= c
FT WO200161009-A2.
XX
PN 23-AUG-2001.
XX
PD 15-FEB-2001; 2001WO-US04828.
XX
PF 15-FEB-2000; 2000US-0182723.
PR 15-FEB-2000; 2000US-0182724.
PR 15-FEB-2000; 2000US-0182733.
PR 22-FEB-2000; 2000US-0183896.
PR 23-FEB-2000; 2000US-0184275.
PR 23-FEB-2000; 2000US-0184482.
PR 23-FEB-2000; 2000US-0184497.
PR 24-FEB-2000; 2000US-0184744.

Db 65 ccccaaaagaagaccagcagctgaaagtcggggtacacctctacacctgggcagcagacagaa 124

QY 258 GAGGATCAGGAGGATCTGCAGGTCACAGGCGGAGCCTGAAGCTCATAGCCAGGAACAG 317

Db 125 gaagatcagatagatcagatcccatgctgcgcagcatggaaggtgatctgcaagagctg 184

QY 318 GGTACCCACAGACTGGGTGTGAAGAT-GGTCTGTATGGCGCAGGAGATGGACCC 376

Db 185 catcagctcaaacaccgggataaatctggttgggtccgcgtcaagtggaagataat 244

QY 377 GCCAATCCACAGAGGTGAACACGCTGAAGAAGTGAACCAATCACAGTGTAAAA 436

Db 245 acctaaagagaaacactgtataatgccagaagcagtggaagagcaaccacaagttaaat 304

QY 437 GAAGACACGTTGAAATGATGAGGCTGCTCTATGTGGAAATTTGTTCAATTAATCT 496

Db 305 gaagacaagtgaacaacgcaagctgttttatattagatatattg--acttaaaactat 361

QY 497 CCAATAAAGCTTACAGCTTCTGCTCAAGAAAAA 534

Db 362 ctcaataaagtttgcagctttccaccaaraaaaaaa 399

RESULT 12

AAH93900 standard; cDNA; 515 BP.

AC AAH93900;

DT 29-JAN-2002 (first entry)

DE Human prostate cDNA sequence #434.

XX Human; prostate cancer; ss: cytostatic; immunostimulant; tumour.

XX Homo sapiens.

XX WO200173032-A2.

PD 04-OCT-2001.

XX 27-MAR-2001; 2001WO-US09919.

PF 27-MAR-2000; 2000US-0536857.

PR 09-MAY-2000; 2000US-0568100.

PR 12-MAY-2000; 2000US-0570737.

PR 13-JUN-2000; 2000US-0593793.

PR 27-JUN-2000; 2000US-0605783.

PR 10-AUG-2000; 2000US-0636215.

PR 29-AUG-2000; 2000US-0651236.

PR 06-SEP-2000; 2000US-0657279.

PR 02-OCT-2000; 2000US-0679426.

PR 10-OCT-2000; 2000US-0685166.

XX (CORI-) CORIXA CORP.

XX Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Kalos MD; Fanger GR, Retter MW, Stolk JA, Day CH, Vedvick TS, Carter D; Li SX, Wang A, Skeiky YAW, Hepler WT, Henderson RA;

DR WPI; 2001-639232/73.

XX New human prostate-specific polypeptides and polynucleotides useful for the diagnosis and treatment of cancer, especially prostate cancer -

XX Claim 1; Page 387; 579pp; English.

XX The invention relates to isolated prostate-specific polynucleotides, polypeptides, fusion proteins of the polypeptides, antibodies raised against the polypeptides (or antigenic epitopes derived from them) and antigen-presenting cells expressing the polypeptides. The antibodies are useful for detecting the presence of cancer, especially prostate cancer. The polypeptides, polynucleotides and

CC the antigen-presenting cells are useful for stimulating and/or expanding T cells specific for a tumour protein, and for inhibiting the development of cancer especially prostate cancer. Compositions comprising the polynucleotide and/or polypeptide are useful for stimulating an immune response, and for treating cancer. The oligonucleotide is useful for detecting cancer. The present sequence is a prostate specific polynucleotide of the invention.

XX Sequence 515 BP; 201 A; 81 C; 112 G; 118 T; 3 other;

QY 187 TGAACACAGCAACACCTGAAGAAGGGGAACCAAGCACTCAACGTCAGGATCCTGCAGCTG 246

Db 141 tgcagcccggtgaatctcagcaagaggaaccaccaactgacaatcagatatgaacctg 200

QY 247 CTCAGGAGGAGAGGATGAGGAGGATCTGCAGGTCGAAGGCCGAGCCCTGAAGCTCAT 306

Db 201 gacaagagagagaa-----ggaacacctcccgatcgaagaacgtaaaagtagaagtgatt 254

QY 307 GCCAGGAACAGGGTCAACCCACAGACTGGGTGTGAGTGTGAAGATGGTCTCATGGCAGG 366

Db 255 gccaggaaatggatcttgaaaagactccggagtgagcgtgagatcggtcttgatgtaaaag 314

QY 367 AGATGGACCCGCCAAATCCAGAGAGGAGGTGAACACGCTGTGAAGAGGTGAAAGCAATCAC 426

Db 315 agaagactccacctaatcctaagcagctgaagactaaagactaaagcagcagagatggcgccat 374

QY 427 AGTGTAAAAGACACACGTTGAAATGATGCA---GGCTGCTCTCTATGTTGGAATTTG- 482

Db 375 aagttaaaaagaagaagctgaagctacacacacatggtgatgtcacattgaaaatgtga 434

QY 483 -----TTCATTAAAAATTCTCCCAATAAAGCTTTACAGCCTTCTGCAAGAAAAA 537

Db 435 ctgaaaaatttgaaaaattctctcaataaagtgttgagtttctctgagaagaaaaaanaa 494

QY 538 A 538

Db 495 a 495

RESULT 13

AAH93807

ID AAH93807 standard; cDNA; 515 BP.

AC AAH93807;

DT 04-OCT-2001 (first entry)

DE Human prostate-specific cDNA sequence P1001C.

XX Human; prostate cancer; prostate-specific; diagnosis; vaccine; cytostatic; gene therapy; metastasis; ss.

XX Homo sapiens.

XX WO200151633-A2.

PD 19-JUL-2001.

XX 16-JAN-2001; 2001WO-US01574.

XX 14-JAN-2000; 2000US-0483672.

XX (CORI-) CORIXA CORP.

XX Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Reed SG; Kalos MD, Fanger GR, Retter MW, Stolk JA, Skeiky YAW; Wang A, Meagher MJ;

XX


```
RESULT 15
AAH02872
ID AAH02872 standard; cDNA; 515 BP.
XX
AC AAH02872;
XX
DT 14-JUN-2001 (first entry)
XX
DE Prostate tumour antigen cDNA sequence for P1001C.
XX
KW Human; prostate tumour antigen; prostate tumour; therapy; diagnosis;
XX prostate cancer; immunogenic; cytostatic; vaccine; ss.
XX
OS Homo sapiens.
XX
PN WO200125272-A2.
XX
PD 12-APR-2001.
XX
PF 04-OCT-2000; 2000WO-US27464.
XX
PR 04-OCT-1999; 99US-0157455.
XX
PA (CORI-) CORIXA CORP.
XX
PI Xu J, Skeiky YAW, Reed SG, Cheever MA;
XX
DR WPI; 2001-245062/25.
XX
CC The present invention describes an isolated polypeptide (I) comprising
CC at least an immunogenic portion of a prostate tumour antigen protein or
CC its variant. (I) have cytostatic activity and can be used in vaccine
CC production. (I), prostate tumour antigen polynucleotides, an antigen
CC presenting cell (APC e.g. a dendritic cell) that expresses (I), and a
CC pharmaceutical composition containing (I) are useful for inhibiting the
CC development of cancer in a patient. Antibodies specific for prostate
CC specific proteins and oligonucleotides that hybridise to a
CC polynucleotide that encodes a prostate specific protein are useful
CC for detecting the presence or absence of a cancer or monitoring the
CC progression the progression of a cancer, especially prostate cancer.
CC AAH02422 to AAH2872, AAH74798 to AAH74821 and AAH74830 are sequences
CC used in the exemplification of the present invention.
XX
SQ Sequence 515 BP; 201 A; 81 C; 112 G; 118 T; 3 other;

Query Match 18.9%; Score 101.6; DB 22; Length 515;
Best Local Similarity 61.2%; Pred. NO. 1.6e-19;
Matches 221; Conservative 0; Mismatches 125; Indels 15; Gaps 3;

QY 187 TGGAAACACGACACCTGTAAGAGGGGACACGACCACTACAGTCAGGATCCTGCAGCTG 246
DB 141 tgcagccggtgaattctcagcaagaggaaacccaaactgaacatcaggattgaaacctg 200
QY 247 CTCAGGAGGGAGAGGATGAGGAGCATCTCAGGTCAAGGGCCGAAGCCTGAAGCTCATA 306
DB 201 gacaagagagagaa-----ggaacacctccgatcgaagaacgtaaagtagaagtgatt 254
QY 307 GCCAGGAACAGGGTCACCCACACACTGGGTGTGAGTGTGAAGATGGTCTCTATGGGCAGG 366
DB 255 gccaggaaatggatctggaaaagactcggagtgagcgtggagatggctctgatgtaaaaag 314
QY 367 AGATGGACCCGCCAATCCAGAGGAGGTGAACACGCCCTGAAGAGGTGAAGCAATCAC 426
DB 315 agaagactccactaatcctaagcatgctaagactaagaagcaggagatggcgagccat 374
QY 427 AGTGTAAAGAGACACCTTGAATGATGCA---GGCTGCTCTATGTGGAAATTTG- 482
```

```
Db 375 aagttaaaaagaagacaagctgaagctacacacatgggtgatgtcacaattgaaaaatgtga 434
QY 483 -----TTCATTAAAAATTCTCCCAATAAAGCTTTTACAGCCTTCTGCAAGAAAAAAA 537
Db 435 ctgaaaaatttgaaaaattctctcaataaagtttgagttttctctggaagaaaaaanaaa 494
QY 538 A 538
Db 495 a 495
```

Search completed: July 1, 2002, 07:11:23
Job time: 1919 sec

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	77.5	17.0	82	1	QBEO	micro glutamic aci
2	74.5	17.5	91	2	S55633	myristylated tequm
3	71	11.0	76	2	C38355	basic proline-rich
4	70	10.8	101	2	B72638	hypothetical prote
5	69.5	10.8	99	2	A5819	nonhistone chromos
6	67.5	10.4	98	2	T02437	hypothetical prote
7	66.5	10.3	79	2	A43732	BR6 protein - mldg
8	66.5	10.3	102	2	A60674	insulin-like growt
9	65	10.1	72	2	S48790	troponin T, cardia
10	62.5	9.7	115	1	FAUTPC	procyelic acidic r
11	62	9.6	111	2	S15073	prothymosin alpha
12	62	9.6	112	1	TNRTA	prothymosin alpha
13	62	9.6	116	2	D72560	hypothetical prote
14	61.5	9.5	102	2	B31512	parathymosin - rat
15	61	9.4	114	2	A30229	procyelic acidic r
16	60.5	9.4	80	2	S54845	GTP cyclohydrolase
17	60.5	9.4	83	2	S78009	aggrecan - pig (fr
18	60.5	9.4	102	2	A32264	parathymosin - hum
19	59.5	9.2	96	2	E70811	hypothetical prote
20	59.5	9.2	109	1	TN8CAL	prothymosin alpha
21	59.5	9.2	110	1	TNHUA	prothymosin alpha
22	59.5	9.2	113	2	T15547	hypothetical prote
23	59.5	9.2	114	1	R5HS2H	Ribosomal protein
24	59.5	9.2	114	2	F84266	50S ribosomal prot
25	59	9.1	80	2	PQ0477	plastil extensin-li
26	59	9.1	81	2	A05231	Balbani ring b ch
27	59	9.1	104	2	C72637	hypothetical prote
28	58.5	9.1	89	2	I54187	small proline-rich
29	58.5	9.1	97	2	S11755	hypothetical prote

euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
 Nature 402, 761-768, 1999
 A:Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.
 A:Reference number: A84420; MUID:20083487
 A:Accession: E84841
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-98 <STO>
 A:Cross-references: GB:AE002093; NID:g3241940; PIDN:AAC23727.1; GSPDB:GNO0139
 C:Genetics:
 A:Gene: T26J13.1; At2g41420
 A:Map position: 2
 A:Introns: 16/1; 84/2

```
Query Match      10.4%; Score 67.5; DB 2; Length 98;  
Best Local Similarity 29.1%; Prod.No. 1.le+02;  
Matches 23; Conservative 10; Mismatches 33; Indels 13; Gaps 3;
```

QY	40	P EEGSPATQRDPAAAQEDEGASACGGPKPEAHSQE-----QGHPQTGCCECDGPDGQE	95
		: : :	
Db	7	PPVGVPFPQGYPP----EGYPKDAYPQGYPPQGYPPQGYPPQGYPPQGYPP----	58
QY	96	M DPNPEEVKTPEEGEKQS	114
		: :	
Db	59	-PPVAPQYPPPPQHQQQKS	76

```

RESULT      7
A43732
BR6 protein - midge (Chironomus tentans) (fragments)
C:Species: Chironomus tentans
C:Date: 03-Mar-1993 #sequence_revision 03-Mar-1993 #text_change 17-Mar-1999
C:Accession: A43732
R:lendahl, U.; Wieslander, L.
Cell 36, 1027-1034, 1984
A:Title: Balbiani ring 6 gene in Chironomus tentans: a diverged member of the Balbiani ring 6 gene family
A:Reference number: A43732; MUID:84156521
A:Accession: A43732
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-79 <LEN>
A:Cross-references: EMBL:K01692

```

	Query Match	10.3%	Score	66.5;	DB	2;	Length	79;
	Best Local Similarity	28.9%;	Pred.	No. le+02;				
	Matches	22;	Conservative	9;	Mismatches	32;	Indels	13; Gaps
QY	12 RPRRYVPEPMIGMRPQEFSDEVEPATPEGEATQRDPDPAAGEGEDGASAGQGPKP	71						
Dd	: :	:	:	: :	:	:	:	:
	12 RPSPERPERPERPER-----PERPEE---REEPERPCKDDE- ---MREKYKR	58						
QY	72 EAHSQEQGHPTQCCEC	87						
Dd	:: :: :	:	:	:	:	:	:	:
	59 RCNDNENRRFDARRKEC	74						

RESULT 8
A60674
insulin-like growth factor I receptor - mouse (fragment)
N:Alternate names: IGF-1 receptor
C:Species: Mus musculus (house mouse)
C:Date: 28-Apr-1993 #sequence_revision 28-Apr-1993 #text_change 23-May-1997
C:Accession: A60674
R:Telford, N.A.; Hogan, A.; Franz, C.R.; Schultz, G.A.
Mol. Reprod. Dev. 27, 81-92, 1990
A:Title: Expression of genes for insulin and insulin-like growth factors an
A:Reference number: A60674; MUID:91064061
A:Accession: A60674
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-102 <TEL>

C; Superfamily: insulin receptor; protein kinase homology
C; Keywords: ATP; growth factor receptor; membrane protein

Query Match	10.3%	Score 66.5	DB 2	Length 102
Best Local Similarity	28.4%	Pred. No. 1.3e+02		
Matches 27	Conservative 13	Mismatches 30	Indels 25	Gaps 6
QY	9	YRPRRYVEPP--EMTGMPEO-----	FSDEVEPATPEGECEPATQDDPAAQ	56
				:
				:
DB	2	YNPK-----IRPSELEIGKIDEMPSQEVSYFYSSEENKPPPEELEMEL--MEPENIG		56
QY	57	EGDEGASAGGPKPEAHSQGGHPQTGCECEDGP	91	
				:
				:
DB	57	EPLDPRESSLPIPERHS---GH-----KAENGP	83	

```

RESULT      9
S48790
troponin T, cardiac muscle (clone HTNT3) - human
C:Species: Homo sapiens (man)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 13-Aug-1999
C:Accession: S48790
R:Farza, H.
submitted to the EMBL Data Library, June 1994
A:Reference number: S48790
A:Accession: S48790
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-72 <FAR>
A:Cross-references: EMBL:X79858; NID:g587429; PIDN:CAA56238.1; PID:g587430
C:Superfamily: troponin T
C:Keywords: cardiac muscle; heart

```

	Query Match	10.1%	Score 65;	DB 2;	Length 72;
Best Local Similarity	34.1%	Pred. No. 1.2e+02;			
Matches 14;	Conservative	6;	Mismatches 21;	Indels	0;
Gaps	0;				

RESULT 10
 FAUTPC
 procyclic acidic repetitive protein precursor (clone pAP3) - Trypanosoma brucei
 N:Alternate names: procyclin
 C:Species: Trypanosoma brucei brucei
 C:Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 22-Jun-1999
 C:Accession: S14896; S06170
 R:Koenig, E.; Delius, H.; Carrington, M.; Williams, R.O.; Roditi, I.
 Nucleic Acids Res. 17, 8727-8739, 1989
 A:Title: Duplication and transcription of procyclin genes in Trypanosoma brucei
 A:Reference number: S14896; MUID:90067841
 A:Accession: S14896
 A:Molecule type: DNA
 A:Residues: 1-115 <R0>
 A:Cross-references: EMBL:X16015; NID:g10512; PIDN:CAA34147.1; PID:g10513
 C:Genetics:

C:Keywords: glycoprotein; tandem repeat
F:1-27/Domain: signal sequence #status predicted <SIG>
F:28-115/Product: procyclic acidic repetitive protein #status predicted <MAT>
F:63-92/Region: 2-residue repeats (E-P)
F:56/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match	9.7%	Score 62.5;	DB 1:	Length 115;
Best Local Similarity	28.0%;	pred. No. 2.9e+02;		
Matches	21:	Conservative	7:	Mismatches 24:
				Indels 23:
				Gaps 3

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 1, 2002, 06:36:39 ; Search time 12.61 Seconds
(without alignments)
356.183 Million cell updates/sec

Title: US-09-782-745-27
Perfect score: 646
Sequence: 1 MSWGRSTYRPRRYVEPP.....DPPNPEVKTPEGEKQSQ 116

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 16058

Minimum DB seq length: 0
Maximum DB seq length: 116

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	646	100.0	116	1 GGE2_HUMAN	Q13066 homo sapien
2	120	18.6	102	1 GGC1_HUMAN	O60829 homo sapien
3	72.5	11.2	85	1 MKKD_RAT	P20468 rattus norv
4	70.5	10.9	96	1 PRP5_HUMAN	P04281 homo sapien
5	62.5	9.7	61	1 PRPE_HUMAN	P02811 homo sapien
6	62.5	9.7	115	1 PARX_TRYBB	P14043 trypanosoma
7	62	9.6	110	1 THYA_MOUSE	P26350 mus musculu
8	62	9.6	111	1 THYA_RAT	P06302 rattus norv
9	61.5	9.5	101	1 THYP_RAT	P04550 rattus norv
10	61	9.4	113	1 MEAL_PIG	Q95313 sus scrofa
11	61	9.4	114	1 PARA_TRYBB	P18764 trypanosoma
12	60.5	9.4	83	1 PGCA_PIG	Q29011 sus scrofa
13	60.5	9.4	101	1 THYP_HUMAN	P20962 homo sapien
14	59.5	9.2	109	1 THYA_BOVIN	P01252 bos taurus
15	59.5	9.2	114	1 RL12_HALHA	P05768 halobacteri
16	58.5	9.1	89	1 CORA_HUMAN	P35321 homo sapien
17	58.5	9.1	109	1 HMG2_HUMAN	P52926 homo sapien
18	58	9.0	99	1 HG14_HUMAN	P05114 homo sapien
19	57.5	8.9	101	1 THSB_BOVIN	P08814 bos taurus
20	57.5	8.9	106	1 VHSB_BPT3	P20322 bacterioph
21	57	8.8	61	1 AERC_AERSO	P09165 aeromonas s
22	57	8.8	97	1 N075_MEDSA	P11728 medicago sa
23	57	8.8	111	1 VG31_BPT4	P17313 bacterioph
24	56.5	8.7	102	1 HG14_CHICK	P12274 gallus gall
25	56.5	8.7	102	1 HSP2_HUMAN	P04554 homo sapien
26	56.5	8.7	110	1 THYA_HUMAN	P06454 homo sapien
27	56	8.7	114	1 ET33_RABIT	P19998 cryptolagus
28	55.5	8.6	102	1 DAPI_HUMAN	P51397 homo sapien
29	55.5	8.6	102	1 HSP2_MACMU	P35297 macaca mula
30	55.5	8.6	103	1 HSP2_MACNE	P35298 macaca neme
31	55	8.5	82	1 MT21_ORYSA	P94029 oryza sativ
32	54.5	8.4	89	1 CORB_HUMAN	P22528 homo sapien
33	54.5	8.4	89	1 CORN_MACMU	P35322 macaca mula

34	54.5	8.4	95	1 HMG2_MOUSE	P17095 mus musculu
35	54.5	8.4	100	1 HG14_BOVIN	P02316 bos taurus
36	54	8.4	104	1 HG15_CHICK	P12902 gallus gall
37	53.5	8.3	84	1 VE4_HPV1A	P06923 human papil
38	53.5	8.3	86	1 CA1C_BOVIN	P25508 bos taurus
39	53.5	8.3	104	1 LEL1_HELAN	P46515 helianthus
40	53.5	8.3	107	1 DFD4_HUMAN	Q9HCU8 homo sapien
41	53	8.2	97	1 CORN_PIG	P35323 sus scrofa
42	53	8.2	105	1 COL1_HSVS7	P25050 herpesvirus
43	52.5	8.1	66	1 IFH2_HIRME	P09944 hirudo medi
44	52.5	8.1	110	1 DCD_HUMAN	P81605 homo sapien
45	52	8.0	65	1 WH11_CANAL	P43074 candida alb

ALIGNMENTS

RESULT 1					
GGE2_HUMAN					
ID GGE2_HUMAN	STANDARD;	PRT;	116 AA.		
AC Q13066;					
DT 15-JUL-1998 (Rel. 36, Created)					
DT 15-JUL-1998 (Rel. 36, Last sequence update)					
DT 30-MAY-2000 (Rel. 39, Last annotation update)					
DE GAGE-2 protein (G antigen 2).					
GN GAGE2.					
OS Homo sapiens (Human).					
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
OX NCBI_TaxID=9606;					
RN [1]					
RP SEQUENCE FROM N.A.					
RC TISSUE=Melanoma;					
RX MEDLINE=95378788; PubMed=7544395;					
RA van den Eynde B., Peeters O., de Backer O., Gaugler B., Lucas S.,					
RA Boon T.;					
RT "A new family of genes coding for an antigen recognized by autologous					
RT cytolytic T lymphocytes on a human melanoma.";					
RL J. Exp. Med. 182:689-698(1995).					
CC -!- FUNCTION: ANTIGEN, RECOGNIZED ON MELANOMA BY AUTOLOGOUS CYTOLYTIC					
CC T LYMPHOCYTES.					
CC -!- TISSUE SPECIFICITY: EXPRESSED IN A VARIETY OF TUMOR TISSUES BUT					
CC NOT IN NORMAL TISSUES, EXCEPT TESTIS.					
CC -!- SIMILARITY: BELONGS TO THE GAGE FAMILY.					
CC -----					
CC This SWISS-PROT entry is copyright. It is produced through a collaboration					
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -					
CC the European Bioinformatics Institute. There are no restrictions on its					
CC use by non-profit institutions as long as its content is in no way					
CC modified and this statement is not removed. Usage by and for commercial					
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/					
CC or send an email to license@isb-sib.ch).					
CC -----					
DR EMBL: U19143; AAA82745.1; -					
DR MIM: 604244; -					
KW Multigene family; Antigen.					
SQ SEQUENCE 116 AA; 12786 MW; DD305D5CA29AF19A CRC64;					

Query Match	100.0%	Score 646;	DB 1;	Length 116;
Best Local Similarity	100.0%	Pred. No. 3.8e-38;		
Matches 116;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY 1 MSWGRSTYRPRRYVEPPPMRPFQSFDEVPATPEEGEPATQRQDPAQAQEGED 60				
Db 1 MSWGRSTYRPRRYVEPPPMRPFQSFDEVPATPEEGEPATQRQDPAQAQEGED 60				
QY 61 EGASAGGPKPEAHSQEGHPQTGCCECDGPDGQEMDPPNPEEVKTPPEGEKQSQ 116				
Db 61 EGASAGGPKPEAHSQEGHPQTGCCECDGPDGQEMDPPNPEEVKTPPEGEKQSQ 116				
RESULT 2				


```
RA Mowatt M.R., Wisdom G.S., Clayton C.E.;
RT "Variation of tandem repeats in the developmentally regulated
   procyclic acidic repetitive proteins of Trypanosoma brucei.";
RL Mol. Cell. Biol. 9:1332-1335(1989).
[2]
RP SEQUENCE FROM N.A.
RX MEDLINE=90258895; PubMed=2342468;
RA Clayton C.E., Fueri J.P., Itzhaki J.E., Bellofatto V., Sherman D.R.,
   Wisdom G.S., Vijayasaraty S., Mowatt M.R.;
RT "Transcription of the procyclic acidic repetitive protein genes of
   Trypanosoma brucei.";
RL Mol. Cell. Biol. 10:3036-3047(1990).
[3]
RP SEQUENCE FROM N.A.
RC STRAIN=427;
RA Vijayasaraty S., Ernest I., Itzhaki J., Sherman D., Mowatt M.R.,
   Michels P.A.M., Clayton C.E.;
RL Submitted (JUN-1992) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: MAJOR SURFACE ANTIGEN OF PROCYCLIC FORMS.
CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
CC -!- DEVELOPMENTAL STAGE: EXPRESSED ONLY AT A CERTAIN STAGE DURING
   DIFFERENTIATION IN THE INSECT VECTOR.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
   between the Swiss Institute of Bioinformatics and the EMBL outstation -
   the European Bioinformatics Institute. There are no restrictions on its
   use by non-profit institutions as long as its content is in no way
   modified and this statement is not removed. Usage by and for commercial
   entities requires a license agreement (See http://www.isb-sib.ch/announce/
   or send an email to license@isb-sib.ch).
-----
DR EMBL; M25787; AAA53283.1; -.
DR EMBL; M33129; AAA30224.1; -.
DR EMBL; X52584; CAA36814.1; -.
DR PIR; A30229; A30229.
DR PIR; S21539; S21539.
KW Signal; Antigen; Repeat; GPI-anchor.
FT SIGNAL 1 27
FT CHAIN 28 92 PROCYCLIC FORM SPECIFIC POLYPEPTIDE A-
   ALPHA.
FT PROPEP 93 114
FT DOMAIN 48 85 6 X 5 AA TANDEM REPEATS OF G-P-E-T.
FT LIPID 92 92 GPI-ANCHOR.
SQ SEQUENCE 114 AA; 11611 MW; FFF2690DAAAE445E CRC64;

Query Match 9.4%; Score 61; DB 1; Length 114;
Best Local Similarity 34.4%; Pred. No. 2.2e+02;
Matches 22; Conservative 8; Mismatches 24; Indels 10; Gaps 4;

Qy 60 DEGAS-----AGQPKPE---AHSQEQGHFOTGCECEDGPGQEMDPNPNPEVKTPPEGE 111
   ||||| :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:|
Db 29 DESASNVIVKGGKREDGPEETGPEETGPE-ETGPEETGPEETGPEET-GPEETE 86
   ||||| :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:|

Qy 112 KQSQ 115
   :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:|
Db 87 PEPE 90

RESULT 12
PGCA_PIG
ID PGCA_PIG STANDARD; PRT; 83 AA.
AC Q29011;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Aggrecan core protein (Cartilage-specific proteoglycan core protein)
DE (CSPCP) (Fragment).
GN AGC1.
OS Sus_scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
```

```
[1]
RN SEQUENCE FROM N.A.
RP TISSUE=Cartilage;
RX MEDLINE=95128522; PubMed=7827755;
RA Barry F.P., Neame P.J., Sasse J., Pearson D.;
RT "Length variation in the keratan sulfate domain of mammalian
   aggrecan.";
RL Matrix Biol. 14:323-328(1994).
CC -!- FUNCTION: THIS PROTEOGLYCAN IS A MAJOR COMPONENT OF EXTRACELLULAR
   MATRIX OF CARTILAGINOUS TISSUES. A MAJOR FUNCTION OF THIS PROTEIN
   IS TO RESIST COMPRESSION IN CARTILAGE. IT BINDS AVIDLY TO
   HYALURONIC ACID VIA AN AMINO-TERMINAL GLOBULAR REGION. MAY PLAY A
   REGULATORY ROLE IN THE MATRIX ASSEMBLY OF THE CARTILAGE.
CC -!- SUBCELLULAR LOCATION: SECRETED; EXTRACELLULAR MATRIX (BY
   SIMILARITY).
CC -!- PTM: CONTAINS MOSTLY CHONDROITIN SULFATE, BUT ALSO KERATAN SULFATE
   CHAINS. N-LINKED AND O-LINKED OLIGOSACCHARIDES (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE AGGREGAN/VERSICAN PROTEOGLYCAN FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
   between the Swiss Institute of Bioinformatics and the EMBL outstation -
   the European Bioinformatics Institute. There are no restrictions on its
   use by non-profit institutions as long as its content is in no way
   modified and this statement is not removed. Usage by and for commercial
   entities requires a license agreement (See http://www.isb-sib.ch/announce/
   or send an email to license@isb-sib.ch).
-----
DR EMBL; S74664; AAC60528.1; -.
DR InterPro; IPR000538; Link.
DR PROSITE; PS01241; LINK; PARTIAL.
DR PROSITE; PS00615; C_TYPE_LLECTIN_1; PARTIAL.
DR PROSITE; PS50041; C_TYPE_LLECTIN_2; PARTIAL.
KW Glycoprotein; Cartilage; Proteoglycan; Repeat.
FT NON_TER 1 1
FT DOMAIN <1 >83 KS.
FT NON_TER 83 83
FT NON_TER 83 83
SQ SEQUENCE 83 AA; 8576 MW; 99DC5641CB0D44A8 CRC64;

Query Match 9.4%; Score 60.5; DB 1; Length 83;
Best Local Similarity 27.3%; Pred. No. 1.8e+02;
Matches 24; Conservative 9; Mismatches 34; Indels 21; Gaps 4;

Qy 24 GPMRPEQFSDEVEPATPEEGEPATQRQDPAAGQEGEDGASAGQGPKEPAHSQEQGHPT 83
   ||||| :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:|
Db 10 GPSATEAPSTSEPPFPSEKPFPS---EPPFPSEPPSEKPSASEPFP-----SEQ--PST 60
   ||||| :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:|

Qy 84 GCECEDGPGQEMDPNPNPEVKTPPEGE 111
   :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:|
Db 61 -----LSAPVPSRTELPGSGE 76

RESULT 13
THYP_HUMAN
ID THYP_HUMAN STANDARD; PRT; 101 AA.
AC P20962;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Parathymosin.
GN PTMS.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
RN SEQUENCE FROM N.A.
RP TISSUE=Kidney;
RX MEDLINE=89149806; PubMed=2537638;
RA Clinton M., Frangou-Lazaridis M., Panneerselvam C., Horecker B.L.;
RT "The sequence of human parathymosin deduced from a cloned human
   kidney cDNA.";
```


Search completed: July 1, 2002, 06:39:48
Job time: 189 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 1, 2002, 06:37:09 ; Search time 21.81 Seconds
(without alignments)
920.101 Million cell updates/sec

Title: US-09-782-745-27
Perfect score: 646
Sequence: 1 MSWRGRSTRPRRYVEPP.....DDPNPEVKTPPEGEKQSQC 116

Scoring table: BLOSUM62

Gapod 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 133496

Minimum DB seq length: 0
Maximum DB seq length: 116

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database :
- 1: sp_archaea.*
 - 2: sp_bacteria.*
 - 3: sp_fungi.*
 - 4: sp_human.*
 - 5: sp_invertebrate.*
 - 6: sp_mammal.*
 - 7: sp_mhc.*
 - 8: sp_organelle.*
 - 9: sp_phage.*
 - 10: sp_plant.*
 - 11: sp_rodent.*
 - 12: sp_virus.*
 - 13: sp_vertebrate.*
 - 14: sp_unclassified.*
 - 15: sp_rviro.*
 - 16: sp_bacteriap.*
 - 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	637	98.6	116	4 Q9UEU5	Q9ueu5 homo sapien
2	229.5	35.5	111	4 Q96GT9	Q96gt9 homo sapien
3	107.5	16.6	110	4 Q96GU1	Q96gu1 homo sapien
4	84.5	13.1	90	5 O18562	O18562 strongyloid
5	82.5	12.8	108	5 Q9GYX8	Q9gyx8 fasciola he
6	79	12.2	90	12 P87914	P87914 human herpe
7	76	11.6	74	11 Q9QZK1	Q9qzk1 cavia porce
8	75	11.6	81	12 Q85063	Q85063 cottontail
9	74.5	11.5	91	12 Q66642	Q66642 equine herp
10	72	11.1	93	4 Q15215	Q15215 homo sapien
11	71.5	11.1	99	5 Q23793	Q23793 chironomus
12	71.5	11.1	107	5 Q9N3B4	Q9n3b4 caenorhabdi
13	71	11.0	75	2 Q9KIL8	Q9kil8 streptomyce
14	71	11.0	114	2 Q9KYF6	Q9kyf6 streptomyce
15	70.5	10.9	100	4 Q9Y3S3	Q9y3s3 homo sapien
16	70.5	10.9	114	4 Q9H1E6	Q9hle6 homo sapien

17	70	10.8	101	17 Q9YEP0	Q9yep0 aeropyrum p
18	69.5	10.8	91	6 Q18928	Q18928 macaca radi
19	69.5	10.8	99	5 Q23794	Q23794 chironomus
20	69	10.7	115	4 Q9ULQ8	Q9ulq8 homo sapien
21	68	10.5	75	2 Q9RK88	Q9rk88 streptomyce
22	68	10.5	82	5 Q9VUQ5	Q9vuq5 drosophila
23	67.5	10.4	98	10 Q80818	Q80818 arabidopsis
24	67	10.4	95	6 Q9GKX4	Q9gkx4 canis famli
25	66.5	10.3	107	4 Q15254	Q15254 homo sapien
26	66	10.2	92	12 Q87076	Q87076 pseudorabie
27	66	10.2	99	5 Q9VPS9	Q9vps9 drosophila
28	65.5	10.1	66	4 Q13902	Q13902 homo sapien
29	64	9.9	82	6 Q9TV59	Q9tv59 canis famli
30	64	9.9	106	4 Q9GZP8	Q9gzp8 homo sapien
31	63.5	9.8	82	6 Q9TV58	Q9tv58 canis famli
32	63.5	9.8	105	4 Q9UED2	Q9ued2 homo sapien
33	63.5	9.8	109	11 Q923P0	Q923p0 mus musculu
34	63.5	9.8	116	6 Q9TV57	Q9tv57 canis famli
35	63	9.8	105	4 Q9NYD3	Q9nyd3 homo sapien
36	63	9.8	107	5 Q9VPS8	Q9vps8 drosophila
37	62.5	9.7	99	6 Q9GKX5	Q9gkx5 canis famli
38	62.5	9.7	101	4 Q9UMZ1	Q9umz1 homo sapien
39	62	9.6	101	11 Q9D0J8	Q9d0j8 mus musculu
40	62	9.6	108	4 Q15202	Q15202 homo sapien
41	62	9.6	110	10 Q9FH00	Q9fhn0 arabidopsis
42	62	9.6	115	12 Q84278	Q84278 canine oral
43	62	9.6	116	17 Q9YB29	Q9yb29 aeropyrum p
44	61.5	9.5	53	13 Q9DFL0	Q9dfd0 gillichthys
45	61.5	9.5	106	13 Q98UF2	Q98uf2 xenopus lae

ALIGNMENTS

RESULT 1

Q9UEU5 PRELIMINARY; PRT; 116 AA.

AC Q9UEU5

DT 01-MAY-2000 (TREMBLrel. 13, Created)

DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE GAGE-8.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A. PubMed=10397259;

RX MEDLINE=93323388; PubMed=10397259;

RA De Backer O., Arden K.C., Boretta M., Vantomme V., De Smet C.,

RA Czekay S., Viars C.S., De Plaen E., Brasseur F., Chomez P.,

RA Van den Eynde B., Boon T., van der Bruggen P.;

RT "Characterization of the GAGE genes that are expressed in various

RT human cancers and in normal testis.

RL Cancer Res. 59:3157-3165(1999).

DR ENBL; AF05473; AAC33676.1; -.

SQ SEQUENCE 116 AA; 12764 MW; DD7052959E66F19A CRC64;

Query Match 98.6%; Score 637; DB 4; Length 116;
Best Local Similarity 99.1%; Pred. No. 5.1e-50;
Matches 115; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSWRGRSTRPRRYVEPPMIGPMRPFQFSDVEPATPEEGPATQRQDPAAQEGED 60
|||||
Db 1 MSWRGRSTRPRRYVEPPMIGPMRPFQFSDVEPATPEEGPATQRQDPAAQEGED 60
|||||
QY 61 EGASAGGPKPEAHSQEQGHQPOTGCECDGPDQGMDDPPNPPEVKTPPEGEKQSQC 116
|||||
Db 61 EGASAGGPKPEAHSQEQGHQPOTGCECDGPDQGMDDPPNPPEVKTPPEGEKQSQC 116
|||||

RESULT 2

DE F58A (HYPOTHETICAL 8.3 KDA PROTEIN).
GN F58A OR SCP8.34C.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2), AND M145;
RA Kormanec J., Sevcikova B., Homerova D.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Saunders D., Harris D.;
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Cerdeno A.M., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RX MEDLINE=97000351; PubMed=8843436;
RA Redenbach M., Kieser H.M., Denapaite D., Eichner A., Cullum J.,
RA Kinashi H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL Mol. Microbiol. 21:77-96(1996).
DR EMBL; AF230490; AAF82062.1; -;
DR EMBL; AL390975; CAC01372.1; -;
SQ SEQUENCE 75 AA; 8263 MW; 392569EF4D29B941 CRC64;

Query Match 11.0%; Score 71; DB 2; Length 75;
Best Local Similarity 27.8%; Pred. No. 15;
Matches 25; Conservative 10; Mismatches 35; Indels 20; Gaps 4;
QY 23 IGPMPPEQFSDEVEPATPEGEPTQRDPAAQGEDEGASAGQGPKEAHSQEQGHQPQ 82
Db 3 VDPTDPTFDDAEAPDPEDRD--VEAPDADAAEQRADVG-----QERDAPL 47

QY 83 TGCCECDGPD-QGEMDPNPEVKPEEGE 111
Db 48 TGVE----PDRADEADLAEQARVPHDEDD 73

RESULT 14
ID Q9KYF6 PRELIMINARY; PRT; 114 AA.
AC Q9KYF6;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE PUTATIVE SECRETED PROTEIN (FRAGMENT).
GN SCC61A.37C.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Oliver K., Harris D.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Bentley S.D., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.

RC STRAIN=A3(2);
RX MEDLINE=97000351; PubMed=8843436;
RA Redenbach M., Kieser H.M., Denapaite D., Eichner A., Cullum J.,
RA Kinashi H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL Mol. Microbiol. 21:77-96(1996).
DR EMBL; AL356595; CAB92282.1; -;
FT NON_TER 1
SQ SEQUENCE 114 AA; 11673 MW; EC396FC15B4B9991 CRC64;

Query Match 11.0%; Score 71; DB 2; Length 114;
Best Local Similarity 29.8%; Pred. No. 23;
Matches 25; Conservative 8; Mismatches 39; Indels 12; Gaps 3;

QY 32 SDEVEPATPEGEPTQRDPAAQGEDEGASAGQGPKEAHSQEQGHPTGCGCEDGP 91
Db 14 NDAEPPGYGEEPPPTTPPG---YGEED-----PPPTTH--EPPPPSGEEPPPPP 61

QY 92 DQEMDPNPEVKTPPEGEKQSQ 115
Db 62 SEEHPTPAPQTEQPPALAEATGSE 85

RESULT 15
ID Q9Y3S3 PRELIMINARY; PRT; 100 AA.
AC Q9Y3S3;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE MUC-B1 PROTEIN (FRAGMENT).
GN MUC-B1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=COLON;
RA Walter A.O., Schwaebler W.J., Thres T., Radecke K., Real F.X.,
RA Dippold W.;
RT "Cloning and characterisation of a novel human mucin gene with 53-
base-pair tandem repeats.";
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ006205; CAB43492.1; -;
FT NON_TER 1
FT NON_TER 100
SQ SEQUENCE 100 AA; 11009 MW; AB58C605152BB911 CRC64;

Query Match 10.9%; Score 70.5; DB 4; Length 100;
Best Local Similarity 29.9%; Pred. No. 22;
Matches 23; Conservative 10; Mismatches 39; Indels 5; Gaps 3;

QY 21 EMIGMPRPEQFSDEVEPA---TFEEGPATQRDPAAQGEDEGASAGQGPKEAHSOE 77
Db 21 EKTKPLRPQQ-QPQCQPAAGTGQRRGSGSSPSADQQAQDREE-AAAAAPTSGRTRTE 78

QY 78 QGHPTGCECEDGPDGQ 94
Db 79 KRKPOQPORRPAAGTGQ 95

Search completed: July 1, 2002, 06:40:24
Job time: 195 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 1, 2002, 06:33:09 ; Search time 23.51 Seconds
(without alignments)
548.046 Million cell updates/sec

Title: US-09-782-745-27

Perfect score: 646

Sequence: 1 MSWRGRSTYRPRRYVEPP.....DPNPPEVKTPGEKQSQ 116

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 11107396 residues

Total number of hits satisfying chosen parameters: 495839

Minimum DB seq length: 0

Maximum DB seq length: 116

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :			
A_Geneseq_032802.*			
1:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*		
2:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*		
3:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*		
4:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*		
5:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*		
6:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*		
7:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*		
8:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*		
9:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*		
10:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*		
11:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*		
12:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*		
13:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*		
14:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*		
15:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*		
16:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*		
17:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*		
18:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*		
19:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*		
20:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*		
21:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*		
22:	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*		

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	646	100.0	116	21	AA1983159
2	638	98.8	116	19	AA1983159
3	400	61.9	76	21	AA1983159
4	271.5	42.0	111	22	AA1983159
5	268.5	41.6	112	22	AA1983159
6	249.5	38.6	106	22	AA1983159
7	229.5	35.5	111	22	AA1983159
8	229.5	35.5	111	22	AA1983159
9	155.5	24.1	60	22	AA1983159
10	141	21.8	115	22	AA1983159
11	120	18.6	102	21	AA1983159

12	120	18.6	102	21	AA1983159	Human secreted pro
13	113.5	17.6	87	21	AA1983159	PAGE1 polypeptide.
14	109.5	17.0	111	22	AA1983159	Human NOV2 protein
15	108.5	16.8	79	21	AA1983159	PAGE3 polypeptide.
16	104	16.1	89	20	AA1983159	Human 5' EST seque
17	96.5	14.9	111	22	AA1983159	Human NOV3 protein
18	96.5	14.9	111	22	AA1983159	Protein encoded by
19	90	13.9	16	20	AA1983159	Tumour rejection a
20	90	13.9	16	20	AA1983159	GAGE tumour reject
21	89.5	13.9	102	22	AA1983159	Peptide #2648 enco
22	89.5	13.9	102	22	AA1983159	Protein #2602 enco
23	89.5	13.9	102	22	AA1983159	Human brain expres
24	89.5	13.9	102	22	AA1983159	Human bone marrow
25	89.5	13.9	102	22	AA1983159	Peptide #2621 enco
26	89.5	13.9	102	22	AA1983159	Peptide #2714 enco
27	89.5	13.9	102	22	AA1983159	Peptide #2599 enco
28	88	13.6	105	18	AA1983159	Collagen-like poly
29	88	13.6	105	18	AA1983159	Silver halide emul
30	88	13.6	105	19	AA1983159	C-terminal of a Sa
31	88	13.6	105	19	AA1983159	Collagen like poly
32	87	13.5	93	18	AA1983159	Collagen-like poly
33	87	13.5	93	18	AA1983159	Silver halide emul
34	87	13.5	93	19	AA1983159	C-terminal of a Sa
35	87	13.5	93	19	AA1983159	Collagen like poly
36	86	13.3	15	17	AA1983159	Tumour rejection a
37	86	13.3	15	19	AA1983159	Antigenic peptide
38	86	13.3	15	19	AA1983159	Antigenic peptide
39	86	13.3	15	20	AA1983159	GAGE tumour reject
40	86	13.3	107	22	AA1983159	Peptide #3676 enco
41	86	13.3	107	22	AA1983159	Peptide #3720 enco
42	86	13.3	107	22	AA1983159	Protein #3590 enco
43	86	13.3	107	22	AA1983159	Human brain expres
44	86	13.3	107	22	AA1983159	Human bone marrow
45	86	13.3	107	22	AA1983159	Peptide #3647 enco

ALIGNMENTS

RESULT 1	
AA1983159	AA1983159 standard; Protein; 116 AA.
ID	AA1983159
XX	AA1983159
AC	AA1983159
DT	24-JUL-2000 (first entry)
DE	GAGE2 polypeptide.
XX	PAGE-4; MAGE; GAGE; reproduction; testis; prostate; fallopian tube;
KW	uterus; placenta; cancer; major histocompatibility complex; MHC; CTL;
KW	cytotoxic T lymphocyte; immune response; antibody; drug delivery;
KW	immunconjugate.
XX	Homo sapiens.
OS	Homo sapiens.
XX	WO200012706-A1
PN	09-MAR-2000.
PD	31-AUG-1999; 99WO-US200046.
PF	01-SEP-1998 98US-0098993.
XX	(US) US DEPT HEALTH & HUMAN SERVICES.
XX	Pastan I, Brinkmann U, Vasmatazis G, Lee B;
XX	WPI; 2000-237869/20.
XX	Isolated PAGE-4 protein and peptide used for inducing a cytotoxic T
XX	lymphocyte response and for raising antibodies which can be used to
XX	detect the presence of PAGE-4 in cell samples or body tissues

DR N-PSDB: AAC02129.

XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for

PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for

PT diagnostic, forensic, gene therapy and chromosome mapping procedures -

XX

XX Claim 13; SEQ ID 6204; 7lpp + CD-ROM; English.

XX

CC The present sequence is a polypeptide encoded by one of a large number

CC of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs

CC were prepared from total human RNAs or polyA+ RNAs derived from 30

CC different tissues. EST sequences usually correspond mainly to the 3'

CC untranslated region (UTR) of the mRNA because they are often obtained

CC from oligo-dT primed cDNA libraries. Such ESTs are not well suited for

CC isolating cDNA sequences derived from the 5' ends of mRNAs and even in

CC those cases where longer cDNA sequences have been obtained, the full 5'

CC UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'

CC ends and can therefore be used to obtain full length cDNAs and genomic

CC DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and

CC chromosome mapping procedures. They are used to obtain upstream

CC regulatory sequences and to design expression and secretion vectors.

XX

XX Sequence 76 AA;

XX

Query Match 61.9%; Score 400; DB 21; Length 76;

Best Local Similarity 97.3%; Pred. No. 4e-30;

Matches 73; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 22 MIGPMRPEQFSEVDEPATPEEGEPATQRDPAAQAQEGDEGASAGQGPKEAHSQEQGHP 81

Db 1 mlgpmrpeqfsevepatpeegpatqxdqpaaaxgdegasagggpkpeahsqeqghp 60

QY 82 QTGCCECDGPDQEM 96

Db 61 qtgcecdgpdqgem 75

RESULT 4

AAM39588

ID AAM39588 standard; Protein; 111 AA.

AC AAM39588;

XX

XX 22-OCT-2001 (first entry)

DE Human polypeptide SEQ ID NO 2733.

XX

XX Human; neurotropic; immunosuppressant; cytostatic; gene therapy; cancer;

KW peripheral nervous system; neuropathy; central nervous system; CNS;

KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;

KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;

KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;

KW leukaemia.

XX

XX Homo sapiens.

XX

XX WO200153312-A1.

PN

XX

XX 26-JUL-2001.

PD

XX

XX 26-DEC-2000; 2000WO-US34263.

PF

XX

XX 21-JAN-2000; 2000US-0489725.

PR

XX 23-APR-2000; 2000US-0552317.

PR

XX 09-JUL-2000; 2000US-0598042.

PR

XX 19-JUL-2000; 2000US-0620312.

PR

XX 03-AUG-2000; 2000US-0653450.

PR

XX 14-SEP-2000; 2000US-0662191.

PR

XX 19-OCT-2000; 2000US-0693036.

PR

XX 29-NOV-2000; 2000US-0727344.

XX

XX (HYSE-) HYSEQ INC.

PA

XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;

PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;

PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;

XX WPI; 2001-442253/47.

DR N-PSDB: AAI58744.

XX

XX Novel nucleic acids and polypeptides, useful for treating disorders

PT such as central nervous system injuries -

XX

PS Example 4; SEQ ID NO 2733; 10078pp; English.

XX

CC The invention relates to human nucleic acids (AAI57798-AAI61369) and

CC the encoded polypeptides (AAM38642-AAM42213) with neurotropic

CC immunosuppressant and cytostatic activity. The polynucleotides are useful

CC in gene therapy. A composition containing a polypeptide or polynucleotide

CC of the invention may be used to treat diseases of the peripheral nervous

CC system, such as peripheral nervous injuries, peripheral neuropathy and

CC localised neuropathies and central nervous system diseases, such as

CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic

CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the

CC utilisation of the activities such as: immune system suppression,

CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic

CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,

CC assays for receptor activity, arthritis and inflammation, leukaemias and

CC C.N.S disorders.

CC Note: The sequence data for this patent did not form part of the printed

CC specification.

XX

XX Sequence 111 AA;

XX

Query Match 42.0%; Score 271.5; DB 22; Length 111;

Best Local Similarity 50.0%; Pred. No. 5.4e-18;

Matches 59; Conservative 13; Mismatches 35; Indels 11; Gaps 3;

QY 1 MSWRGRSTYRPRRYVEPEPMIGPMRPEQFSDEVEPA--TPEEGEPATQRDPAAQAQ- 57

Db 1 miwgrstyrprtrsvpppeligpm-----lepgdeapqgeepptesrdpaggqr 52

QY 58 GEDEGASAGQGPKEAHSQEQGHPQTCCECDGPDQEMDPNPNEVKTPEEGEKQSQ 115

Db 53 eedggaatqvpdleadlqelsqsktggecgngpddqgkllpkseqfkmpggdrpq 110

RESULT 5

ABG05297

ID ABG05297 standard; Protein; 112 AA.

XX

AC ABG05297;

XX

XX 13-FEB-2002 (first entry)

DT

XX

XX Novel human diagnostic protein #5288.

DE

XX

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;

KW food supplement; medical imaging; diagnostic; genetic disorder.

KW

XX

XX Homo sapiens.

OS

XX

XX WO200175067-A2.

PN

XX

XX 11-OCT-2001.

PD

XX

XX 30-MAR-2001; 2001WO-US08631.

PF

XX

XX 31-MAR-2000; 2000US-0540217.

PR

XX 23-AUG-2000; 2000US-0649167.

PR

XX

XX (HYSE-) HYSEQ INC.

PA

XX

XX Drmanac RT, Liu C, Tang YT;

PI

```
XX WPI; 2001-6393362/73.
DR N-PSDB; AAS69484.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
XX Claim 20; SEQ ID NO 35656; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 112 AA;
SQ
Query Match 41.6%; Score 268.5; DB 22; Length 112;
Best Local Similarity 49.2%; Pred. No. 1e-17;
Matches 58; Conservative 14; Mismatches 35; Indels 11; Gaps 3;
Qy 1 MSWRGSTRPRPRRYVEPPMIGMPMRPEQSFDEVEPA--TPEEGEPATQRDPAAAOE- 57
Db | ||||| ||||| | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| |
Qy 58 GEDEGASAGGPKPEAHSQEGHPQTGCECEDGPDGQEMDPNPNBEVKTPEEGEKOSQ 115
Db ||::: | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 54 kedqgsaktqvpdleadlqlsqsgktgcecgngpddqgkllpkseqfkmpcgddrqpq 111
Db ||::: | | | | | | | | | | | | | | | | | | | | | | | | | | |
RESULT 6
AAM41374
ID AAM41374 standard; Protein; 106 AA.
XX
XX AAM41374;
AC
XX
XX 22-OCT-2001 (first entry)
DT
XX
XX Human polypeptide SEQ ID NO 6305.
DE
XX
XX Human; neotropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.
XX
XX Homo sapiens.
OS
XX
XX WO200153312-A1.
PN
XX
XX 26*JUL-2001.
PD
XX
XX 26-DEC-2000; 2000WO-US34263.
PF
```

```
XX 21-JAN-2000; 2000US-0488725.
DR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX
XX (HYSE-) HYSEQ INC.
PA
XX
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
XX WPI; 2001-442253/47.
DR N-PSDB; AAI60530.
XX
XX Novel nucleic acids and polypeptides, useful for treating disorders
XX such as central nervous system injuries -
PT
XX
XX Example 2; SEQ ID NO 6305; 10078pp; English.
PS
XX
XX The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemia and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
XX Sequence 106 AA;
SQ
Query Match 38.6%; Score 249.5; DB 22; Length 106;
Best Local Similarity 50.5%; Pred. No. 5.7e-16;
Matches 56; Conservative 10; Mismatches 34; Indels 11; Gaps 3;
Qy 1 MSWRGSTRPRPRRYVEPPMIGMPMRPEQSFDEVEPA--TPEEGEPATQRDPAAAOE- 57
Db | ||||| ||||| | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| |
Qy 2 miwrgstryrprsrsvppellgpm-----lepdeeppqgeepptesrdpapgqr 53
Db ||::: | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 58 GEDEGASAGGPKPEAHSQEGHPQTGCECEDGPDGQEMDPNPNBEVKTPE 108
Db ||::: | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 54 eedqgaetqvpdleadlqlsqsgktgdecgdpdvqgkllkseqfkmpke 104
Db ||::: | | | | | | | | | | | | | | | | | | | | | | | | | | |
RESULT 7
AAM78785
ID AAM78785 standard; Protein; 111 AA.
XX
XX AAM78785;
AC
XX
XX 06-NOV-2001 (first entry)
DT
XX
XX Human protein SEQ ID NO 1447.
DE
XX
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation.
XX
XX Homo sapiens.
OS
```


ID ABG05299 standard; Protein; 60 AA.
XX AAE08583;
AC AAE08583 standard; Protein; 115 AA.
DT 13-FEB-2002 (first entry)
XX AAE08583;
DE 01-NOV-2001 (first entry)
XX Human NOV4 protein.
XX Human; NOVX; G-antigen; GAGE-like protein; interferon;
KW G-protein coupled receptor; GPCR; hepatocyte nuclear factor;
KW mast cell protease; gene therapy; proliferative disorder; cancer;
XX immune disorder; hepatic disorder; cirrhosis; viral infection;
XX hepatitis; neuroolfactory system-related disorder; neurological disorder;
XX Parkinson's disease; infertility; autoimmune disease; arthritis;
XX multiple sclerosis; allergy; wound healing; cytostatic; nootropic;
XX immunosuppressive; neuroprotective; vulnerary; hepatotropic.
OS Homo sapiens.
XX WO200175067-A2.
XX 11-OCT-2001.
XX 30-MAR-2001; 2001WO-US08631.
XX 31-MAR-2000; 2000US-0540217.
XX 23-AUG-2000; 2000US-0649167.
XX (HYSE-) HYSEQ INC.
XX Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
XX N-PSDB; AAS69486.
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX Claim 20; SEQ ID No 35658; 103pp; English.
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX Sequence 60 AA;
SQ

Query Match 24.1%; Score 155.5; DB 22; Length 60;
Best Local Similarity 50.0%; Pred. No. 1.7e-07;
Matches 34; Conservative 7; Mismatches 14; Indels 13; Gaps 3;
QY 1 MSWRGRSYRPRRVEPEMIGMPRPEQFSDEVEPA--TPEEGEPATQRQDPAAQEG 58
Db 2 mlwrgstprprrrsvppblgpm-----lepgdeeqqeeptcsrdpa---pg 50
QY 59 EDEGASAG 66
Db 51 sgerrrsg 58

RESULT 10
AAE08583
ID AAE08583 standard; Protein; 115 AA.
XX AAE08583;
AC AAE08583;
DT 01-NOV-2001 (first entry)
XX Human NOV4 protein.
XX Human; NOVX; G-antigen; GAGE-like protein; interferon;
KW G-protein coupled receptor; GPCR; hepatocyte nuclear factor;
KW mast cell protease; gene therapy; proliferative disorder; cancer;
XX immune disorder; hepatic disorder; cirrhosis; viral infection;
XX hepatitis; neuroolfactory system-related disorder; neurological disorder;
XX Parkinson's disease; infertility; autoimmune disease; arthritis;
XX multiple sclerosis; allergy; wound healing; cytostatic; nootropic;
XX immunosuppressive; neuroprotective; vulnerary; hepatotropic.
OS Homo sapiens.
XX WO200161009-A2.
XX 23-AUG-2001.
XX 15-FEB-2001; 2001WO-US04828.
XX 15-FEB-2000; 2000US-0182723.
XX 15-FEB-2000; 2000US-0182724.
XX 15-FEB-2000; 2000US-0182733.
XX 22-FEB-2000; 2000US-0183896.
XX 23-FEB-2000; 2000US-0184275.
XX 23-FEB-2000; 2000US-0184482.
XX 23-FEB-2000; 2000US-0184497.
XX 24-FEB-2000; 2000US-0184744.
XX 13-APR-2000; 2000US-0197083.
XX 10-AUG-2000; 2000US-0224157.
XX 18-SEP-2000; 2000US-0233405.
XX 27-SEP-2000; 2000US-0236060.
XX 02-JAN-2001; 2001US-0259414.
XX 18-JAN-2001; 2001US-0262454.
XX 14-FEB-2001; 2001US-0783429.
XX (CURA-) CURAGEN CORP.
XX Malvankar UM, Tchernev VT, Padigaru M, Taupier RJ, Spytek KA;
PI Majumder K, Guo X, Spaderna SK, Boldog FL;
PI N-PSDB; AAD14983.
XX WPI; 2001-514775/56.
XX Isolated novel polypeptides useful for diagnosis of and treating
PT cancer, infertility, autoimmune diseases, arthritis, multiple
PT sclerosis, allergies, wound healing and hepatic disorders -
XX Claim 1; Page 14; 140pp; English.
XX The present sequence is a human NOV4 protein. The NOVX protein has
CC homology with one of G-antigen (GAGE)-like protein, interferon,
CC G-protein coupled receptor (GPCR), hepatocyte nuclear factor or mast
CC cell protease. The NOVX is useful for treating or preventing a pathology
CC associated with NOVX. It is also useful for determining the presence or
CC amount of NOVX DNA in a sample, for identifying a potential therapeutic
CC agent and in gene therapy. It is also useful for determining the presence
CC of or predisposition to a disease associated with altered levels of NOVX.
CC It is also useful for the diagnosis and treatment of proliferative
CC disorders, e.g., cancer, immune disorders, hepatic disorders, e.g.,
CC cirrhosis, viral infections, e.g., hepatitis, neuroolfactory
CC system-related disorders, neurological disorders, e.g., Parkinson's
CC disease, infertility, autoimmune diseases, arthritis, multiple sclerosis,
CC allergies and wound healing.
XX

115 AA;

21.8%; Score 141; DB 22; Length 115;
51.9%; Pred. No. 7.6e-06;
ative 9; Mismatches 7; Indels 1

YYVEPPMIGMRPEQFSDEVPA--TPEEGEPATQRQDP 52
:
:::|:| | :|| |:| |:|:|
slqppelllgam-----leptdeepkeekpptksrnp 46

RESULT 11

protein; 102 AA.

AA
AC AAY83158:

DT 24-JUL-2000 (first entry)

DE PAGE-4 polypeptide.

reproduction; testis; prostate; fallopian tube; cancer; major histocompatibility complex; MHC; CTL; te; immune response; antibody; drug delivery;

OS Homo sapiens.

PN WO200012706-A1.

09-MAR-2000.

31-AUG-1999: 99WO-US20046.

PR 01-SEP-1998; 98US-00989993.

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

PI Pastan I, Brinkmann U, Vasmatazis G, Lee B;

DR WPI; 2000-237869/20.

XX
XX
XXXXXXXXXXXXX

and for raising antibodies which can be used to
of PAGE-4 in cell samples or body tissues

PS Disclosure; Figure 1; 63pp; English.

ferentially expressed in normal male and female e.g. prostate, testis, fallopian tube, uterus and ovary. It is also over expressed in prostate cancer, testicular cancer and endometrial cancer. The expression pattern makes it a target for vaccine based therapy of such neoplasms.

peptide which induces a cytotoxic T_H1 response when bound to a major histocompatibility complex (MHC) molecule or the isolated PAGE-4 protein can be used in conjunction with a cytotoxic T lymphocyte response inducing PAGE-4 including cancer cells of the testis. The nucleic acids encoding PAGE-4 or fragments can also be used in these compositions. PAGE-4 and its peptide fragments can be used in conjunction with PAGE-4 and PAGE-4 coding sequences in cell lines. The presence of PAGE-4 in tissues which are under investigation can be indicative of the spread of cancer. PAGE-4 can also be used to raise antibodies. PAGE-4 can also be used as the targeting group of antibodies. PAGE-4 can also be used as the targeting group of immunising toxins used in therapeutic applications. PAGE-4 can also be used as the targeting group of toxins for drug delivery systems.

XX	Sequence	102 AA:
SO		

Query Match	18.6%;	Score 120;	DB 21;	Length 102;
Best Local Similarity	34.5%;	Pred. No. 0.0006;		
Matches 39;	Conservative 10;	Mismatches 52;	Indels 1	

```

yy 1 MSWRGRSTYRPRRRYVEPPEMIGMPEQFSEVEPATPEEGEPATQRQDPAAAAQEGED 60
    || || || || || || || || || || || || || || || || || || || || ||
bb 1 msarvrsrqrdqgq-eapdvva-----fvapsgqqeeptdndtpeqq--er 48

```

61 EGASAGQGPKPEAHSQEQGHQGTGCECEDGPDQEQMDPPNPPEEVKTPPEGEKQ 113
 49 eqtpoleerkveqdcqemdlekrserqgdsdvektppnpkhaktkeagdgq 101

RESULT 12

AY52998
AY52998 standard: Protein: 102 AA

AA52998:

29-FEB-2000 (first entry)

Human secreted protein clone bn365_53 protein sequence SEQ ID NO:2.

Human; secreted protein; nutritional; cytokine; cell proliferation; differentiation; immune stimulating; vaccine; suppression; haematopoiesis regulation; tissue growth; activin; inhibin; chemotactic; chemokinetic; haemostatic; thrombolytic; receptor; ligand; anti-inflammatory; cadherin; tumour invasion suppressor; tumour inhibition; gene therapy

Homo sapiens.

WO9957132-A1.

11-NOV-1999.

07-MAY-1999; 99WO-US09970.

07-MAY-1998; 98US-0084564.

22-JUL-1998; 98US-0093712.

10-AUG-1998; 98US-0095880.

06-MAY-1999; 99US-0096068.

(GEMY) GENETICS INST INC.

Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;
Merberg D, Treacy M, Agostino MJ, Steininger RJ, Bowman MR;
DiBlasio-Smith E, Widom A;

WPI: 2000-052937/04.

DR N-PSDB; AAZ33316.

New polynucleotides encoding secreted human proteins, derived from adult placenta, adult retina, fetal brain, fetal

Claim 9; Page 360-361; 492pp; English.

The present invention describes new human secreted proteins which were isolated from adult placenta, adult retina, foetal brain, foetal kidney adult blood, adult brain, adult thyroid, adult bladder, adult neural tissue, adult testes, and adult lymph node cDNA libraries. The human secreted proteins, and the polynucleotide encoding them, are predicted to have biological activities which would make them suitable for treating, preventing or ameliorating medical conditions in humans and animals. Suggested activities include nutritional activity, cytokine and cell proliferation/differentiation activity, immune stimulating (e.g. as vaccines) or suppressing activity, haematopoiesis regulating activity, tissue growth activity, activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic and thrombolytic

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 1, 2002, 06:11:45 ; Search time 1773.93 seconds
(without alignments)
6346.630 Million cell updates/sec

Title: US-09-782-745-14
Perfect score: 538
Sequence: 1 ACGCAGGGAGCTGTAGGC.....CTGCAAGAGAAAAA 538

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- GenEmbl:*
- 1: gb_ba:*
 - 2: gb_htg:*
 - 3: gb_in:*
 - 4: gb_om:*
 - 5: gb_ov:*
 - 6: gb_pat:*
 - 7: gb_ph:*
 - 8: gb_pl:*
 - 9: gb_pr:*
 - 10: gb_ro:*
 - 11: gb_sts:*
 - 12: gb_sy:*
 - 13: gb_un:*
 - 14: gb_vi:*
 - 15: em_ba:*
 - 16: em_fun:*
 - 17: em_hum:*
 - 18: em_in:*
 - 19: em_mu:*
 - 20: em_om:*
 - 21: em_or:*
 - 22: em_ov:*
 - 23: em_pat:*
 - 24: em_ph:*
 - 25: em_pl:*
 - 26: em_ro:*
 - 27: em_sts:*
 - 28: em_un:*
 - 29: em_vi:*
 - 30: em_htg_hum:*
 - 31: em_htg_inv:*
 - 32: em_htg_other:*
 - 33: em_htg_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description
-----	-----	-----	-----	-----	-----	-----	-----

1	538	100.0	538	6	AR028488	AR028488 Sequence
2	538	100.0	538	6	155851	155851 Sequence 14
3	533.2	99.1	551	9	BC018052	BC018052 Homo sapi
4	531.6	98.8	1245	6	AX285022	AX285022 Sequence
5	530	98.5	530	9	HSU19143	U19143 Human GAGE-
6	525.4	97.7	528	9	AF055473	AF055473 Homo sapi
7	512.8	95.3	540	6	AR028490	AR028490 Sequence 16
8	512.8	95.3	540	6	155853	155853 Sequence 16
9	510.2	94.8	539	6	AR028492	AR028492 Sequence 18
10	510.2	94.8	539	6	155855	155855 Sequence 18
11	504.8	93.8	532	6	AR028491	AR028491 Sequence 17
12	504.8	93.8	532	6	155854	155854 Sequence 17
13	501.4	93.2	528	6	AX334151	AX334151 Sequence
14	501.4	93.2	528	9	HSU19145	U19145 Human GAGE-
15	498.8	92.7	527	9	HSU19147	U19147 Human GAGE-
16	497.6	92.5	524	9	AF058988	AF058988 Homo sapi
17	496.8	92.3	524	9	HSU19146	U19146 Human GAGE-
18	488	90.7	526	9	AF055474	AF055474 Homo sapi
19	429.2	79.8	560	6	AR028489	AR028489 Sequence 15
20	429.2	79.8	560	6	155852	155852 Sequence 15
21	421.2	78.3	552	9	HSU19144	U19144 Human GAGE-
22	374.4	69.6	646	6	AR028482	AR028482 Sequence 1
23	374.4	69.6	646	6	155845	155845 Sequence 1
24	374.4	69.6	646	9	HSU19142	U19142 Human GAGE-
25	374.4	69.6	648	6	157317	157317 Sequence 1
c 26 301.4 56.0 530 6 AX284300						
c 27 299.2 55.6 365 6 AX284692						
28	209.6	39.0	662	9	BC004861	BC004861 Homo sapi
29	209.6	39.0	676	9	AF058989	AF058989 Homo sapi
30	185.8	34.5	648	9	BC009232	BC009232 Homo sapi
31	182.4	33.9	493	9	HSA318881	HS318881 Homo sapi
32	179.8	33.4	620	9	HS318880	HS318880 Homo sapi
33	173.8	32.3	611	6	AX226501	AX226501 Sequence
c 34 163.2 30.3 79539 9 AC093664						
c 35 163.2 30.3 240000 2 AC009528						
c 36 156 29.0 580 6 AX078298						
37	146.8	27.3	763	9	BC009230	BC009230 Homo sapi
c 38 140.8 26.2 20587 9 HSL18566A						
c 39 140.8 26.2 62493 9 HSL193G15						
c 40 140.8 26.2 62493 9 HSL193G15						
c 41 140.8 26.2 185713 2 AL645949						
c 42 131.4 24.4 9531 9 AF055475						
c 43 131.4 24.4 48802 2 AC068431						
c 44 130.4 24.2 475 6 AX226497						
c 45 129.8 24.1 194418 9 AF235098						

ALIGNMENTS

RESULT 1	AR028488	AR028488	538 bp	DNA	linear	PAT 29-SEP-1999
LOCUS	Sequence 14	from patent US 5858689.				
DEFINITION	Sequence 14	from patent US 5858689.				
ACCESSION	AR028488					
VERSION	AR028488.1	GI:5940461				
KEYWORDS	Unknown.					
SOURCE	Unknown.					
ORGANISM	Unclassified.					
REFERENCE	1 (bases 1 to 538)					
AUTHORS	van der Bruggen, P., van den Eynde, B., DeBacker, O. and Boon-Falleur, T.					
TITLE	Isolated peptides derived from the gage tumor rejection antigen precursor and uses thereof					
JOURNAL	Patent: US 5858689-A 14 12-JAN-1999;					
FEATURES	Location/Qualifiers					
Source	1..538					
BASE COUNT	160 a	116 c	155 g	107 t		
ORIGIN						

Query Match		100.0%; Score 538; DB 6; Length 538;	
Best Local Similarity		100.0%; Pred. No. 6.6e-120;	
Matches 538; Conservative		0; Mismatches 0; Indels 0; Gaps 0;	
QY	1	ACGCCAGGAGCTGTGAGGAGTGTGTTGGTTCCTGCCGTCGCGACTCTTTTTCCTCT	60
Db	1	ACGCCAGGAGCTGTGAGGAGTGTGTTGGTTCCTGCCGTCGCGACTCTTTTTCCTCT	60
QY	61	ACTGAGATTCATCTGTGTGAATATGAGTTGGCGAGGAGATCGACCTATCGGCCTAGAC	120
Db	61	ACTGAGATTCATCTGTGTGAATATGAGTTGGCGAGGAGATCGACCTATCGGCCTAGAC	120
QY	121	CAAGAGCTACGTAGAGCTCCTGAAATGATTGGCCCTATGCGGCCCGAGCAGTTCAGTG	180
Db	121	CAAGAGCTACGTAGAGCTCCTGAAATGATTGGCCCTATGCGGCCCGAGCAGTTCAGTG	180
QY	181	ATGAAGTGGACCAACAGCACTTGAAGAGGGGAACCCAGCAACTCAACGTCAGGATCCTG	240
Db	181	ATGAAGTGGACCAACAGCACTTGAAGAGGGGAACCCAGCAACTCAACGTCAGGATCCTG	240
QY	241	CAGCTGCTCAGGAGGAGAGGATGAGGAGGATCTGCAGGTCAAGGCCCGAAGCCTGAAG	300
Db	241	CAGCTGCTCAGGAGGAGAGGATGAGGAGGATCTGCAGGTCAAGGCCCGAAGCCTGAAG	300
QY	301	CTCATAGCCAGAACAGGGTCAACCCACAGACTGGTGTGAGTGTGAAGATGTTGGAAT	360
Db	301	CTCATAGCCAGAACAGGGTCAACCCACAGACTGGTGTGAGTGTGAAGATGTTGGAAT	360
QY	361	GGCAGGAGATGGACCCGCCAAATCCAGAGGAGTGAACCGCTGAAGAGGTGAAAGC	420
Db	361	GGCAGGAGATGGACCCGCCAAATCCAGAGGAGTGAACCGCTGAAGAGGTGAAAGC	420
QY	421	AATCAGAGTGTAAAGAGACACGTTGAAATGATGCGAGGTGCTCTATGTTGGAAT	480
Db	421	AATCAGAGTGTAAAGAGACACGTTGAAATGATGCGAGGTGCTCTATGTTGGAAT	480
QY	481	TGTTCAATTAATCTCCCAATAAGCTTTACAGCTTCTGCAAGAGTGAAGAGC	538
Db	481	TGTTCAATTAATCTCCCAATAAGCTTTACAGCTTCTGCAAGAGTGAAGAGC	538
RESULT 2		538 bp DNA linear PAT 07-OCT-1997	
LOCUS I55851		Sequence 14 from patent US 5648226.	
DEFINITION		Isolated peptides derived from tumor rejection antigens, and their	
ACCESSION		I55851	
VERSION		I55851.1 GI:2476645	
KEYWORDS		Unknown.	
SOURCE		Unknown.	
ORGANISM		Unclassified.	
REFERENCE		1 (bases 1 to 538)	
AUTHORS		Van den Eynde,B., DeBacker,O. and Boon-Falleur,T.	
TITLE		Isolated peptides derived from tumor rejection antigens, and their	
JOURNAL		Patent: US 5648226-A 14 15-JUL-1997;	
FEATURES		Location/Qualifiers	
source		I..538	
BASE COUNT		160 a 116 c 155 g 107 t	
ORIGIN		100.0%; Score 538; DB 6; Length 538;	
Best Local Similarity		100.0%; Pred. No. 6.6e-120;	
Matches 538; Conservative		0; Mismatches 0; Indels 0; Gaps 0;	
QY	1	ACGCCAGGAGCTGTGAGGAGTGTGTTGGTTCCTGCCGTCGCGACTCTTTTTCCTCT	60
Db	1	ACGCCAGGAGCTGTGAGGAGTGTGTTGGTTCCTGCCGTCGCGACTCTTTTTCCTCT	60
QY	61	ACTGAGATTCATCTGTGTGAATATGAGTTGGCGAGGAGATCGACCTATCGGCCTAGAC	120

Db	61	ACTGAGATTCATCTGTGTGAATATGAGTTGGCGAGGAGATCGACCTATCGGCCTAGAC	120
QY	121	CAAGAGCTACGTAGAGCTCCTGAAATGATTGGCCCTATGCGGCCCGAGCAGTTCAGTG	180
Db	121	CAAGAGCTACGTAGAGCTCCTGAAATGATTGGCCCTATGCGGCCCGAGCAGTTCAGTG	180
QY	181	ATGAAGTGGACCAACAGCACTTGAAGAGGGGAACCCAGCAACTCAACGTCAGGATCCTG	240
Db	181	ATGAAGTGGACCAACAGCACTTGAAGAGGGGAACCCAGCAACTCAACGTCAGGATCCTG	240
QY	241	CAGCTGCTCAGGAGGAGAGGATGAGGAGGATCTGCAGGTCAAGGCCCGAAGCCTGAAG	300
Db	241	CAGCTGCTCAGGAGGAGAGGATGAGGAGGATCTGCAGGTCAAGGCCCGAAGCCTGAAG	300
QY	301	CTCATAGCCAGAACAGGGTCAACCCACAGACTGGTGTGAGTGTGAAGATGTTGGAAT	360
Db	301	CTCATAGCCAGAACAGGGTCAACCCACAGACTGGTGTGAGTGTGAAGATGTTGGAAT	360
QY	361	GGCAGGAGATGGACCCGCCAAATCCAGAGGAGTGAACCGCTGAAGAGGTGAAAGC	420
Db	361	GGCAGGAGATGGACCCGCCAAATCCAGAGGAGTGAACCGCTGAAGAGGTGAAAGC	420
QY	421	AATCAGAGTGTAAAGAGACACGTTGAAATGATGCGAGGTGCTCTATGTTGGAAT	480
Db	421	AATCAGAGTGTAAAGAGACACGTTGAAATGATGCGAGGTGCTCTATGTTGGAAT	480
QY	481	TGTTCAATTAATCTCCCAATAAGCTTTACAGCTTCTGCAAGAGTGAAGAGC	538
Db	481	TGTTCAATTAATCTCCCAATAAGCTTTACAGCTTCTGCAAGAGTGAAGAGC	538
RESULT 3		551 bp mRNA linear PRI 06-DEC-2001	
LOCUS BC018052		Homo sapiens, G antigen 8, clone MGC:26395 IMAGE:4812462, mrna,	
DEFINITION		complete cds.	
ACCESSION		BC018052	
VERSION		BC018052.1 GI:17390105	
KEYWORDS		MGC.	
SOURCE		human.	
ORGANISM		Homo sapiens	
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
AUTHORS		Mammalia; Euthera; Primates; Catarrhini; Homiidae; Homo.	
TITLE		1 (bases 1 to 551)	
JOURNAL		Strausberg,R.	
REMARK		Direct Submission	
COMMENT		Submitted (03-DEC-2001), National Institutes of Health, Mammalian	
		Gene Collection (MGC), Cancer Genomics Office, National Cancer	
		Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,	
		USA	
		NIH-MGC Project URL: http://mgc.nci.nih.gov	
		Contact: MGC help desk	
		Email: cgaphs-r@mail.nih.gov	
		Tissue Procurement: Miklos Palkovits, M.D., Ph.D.	
		cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki	
		Toshiyuki and Piero Carninci (RIKEN)	
		cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)	
		DNA Sequencing by: Sequencing Group at the Stanford Human Genome	
		Center, Stanford University School of Medicine, Stanford, CA 94305	
		Web site: http://www-shgc.stanford.edu	
		Contact: (Dickson, Mark) mcdpaxil.stanford.edu	
		Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,	
		R. M.	
		Clone distribution: MGC clone distribution information can be found	
		through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov	
		Series: IRAC Plate: 32 Row: K Column: 1	
		This clone was selected for full length sequencing because it	
		passed the following selection criteria: matched mRNA gi: 4503878.	
FEATURES		Location/Qualifiers	
source		1..551	
		/organism="Homo sapiens"	
		/db_xref="LocusID:26749"	

JOURNAL Submitted (28-DEC-1994) Benoit J Van Den Eynde, Ludwig Institute
For Cancer Research, 74 Avenue Hippocrate, BRUSSELS, 1200, BELGIUM

FEATURES
source 1..530
/organism="Homo sapiens"
/db_xref="taxon:9606"
/sex="female"
/cell_line="M22-MEL.43"
/tissue_type="melanoma"
/dev_stage="adult"
84..434
/codon_start=1
/product="GAGE-2 protein"
/protein_id="AAA82745.1"
/db_xref="GI:914901"
/translation="MSWRGRSTYRPRRPRRYVEPPMIGMRPEQFSDVEVPATPEEGE
PATORDPAAOEGEDGASAGQGPKEAHSQEQHPQTGCECEDGPDQEMDPNPE
EVKTPPEGEKQSQ"

CDS
152 a 116 c 155 g 107 t

Query Match 98.5%; Score 530; DB 9; Length 530;
Best Local Similarity 100.0%; Pred. No. 5.7e-118;
Matches 530; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACGCCAGGAGCTGTGAGCAGTCTGTGTGGTTCTCGCGTCCGGACTCTTTTCTCT 60
Db 1 ACGCCAGGAGCTGTGAGCAGTCTGTGTGGTTCTCGCGTCCGGACTCTTTTCTCT 60

Qy 61 ACTGAGATTCATCTGTGTGAATATATGAGTTGGCGAGGAAGATGACCTATCGGCCTAGAC 120
Db 61 ACTGAGATTCATCTGTGTGAATATATGAGTTGGCGAGGAAGATGACCTATCGGCCTAGAC 120

Qy 121 CAAGACGCTACGTAGAGCTCTCTGAATGATGTTGGGCTATGCGGCCCGAGAGTTCAAGT 180
Db 121 CAAGACGCTACGTAGAGCTCTCTGAATGATGTTGGGCTATGCGGCCCGAGAGTTCAAGT 180

Qy 181 ATGAAGTGAACAGCAACACCTCTGAAGAGGGGAACAGCAACTCAACGTCAGGATCCTG 240
Db 181 ATGAAGTGAACAGCAACACCTCTGAAGAGGGGAACAGCAACTCAACGTCAGGATCCTG 240

Qy 241 CAGCTGCTCAGGAGGAGGATGAGGAGCATCTCGAGTCAAGGCCGAGAGCTGAAG 300
Db 241 CAGCTGCTCAGGAGGAGGATGAGGAGCATCTCGAGTCAAGGCCGAGAGCTGAAG 300

Qy 301 CTATAGCAGCAAGAGGTCACCCACAGACTGGGTGTGAGTGTGAAGATGGTCTGATG 360
Db 301 CTATAGCAGCAAGAGGTCACCCACAGACTGGGTGTGAGTGTGAAGATGGTCTGATG 360

Qy 361 GGCAGGATGAGACCGCCCAATCCAGAGAGTGAACGCCCTGAAGAGGTGAAGAGC 420
Db 361 GGCAGGATGAGACCGCCCAATCCAGAGAGTGAACGCCCTGAAGAGGTGAAGAGC 420

Qy 421 AATCAGTGTAAAAGACACACGTTGAAATGATGACAGGCTGCTCTATGTTGGAATT 480
Db 421 AATCAGTGTAAAAGACACACGTTGAAATGATGACAGGCTGCTCTATGTTGGAATT 480

Qy 481 TGTTCAATTAATTTCTCCCAATAAAGCTTTTACAGCCTTTTTCGAAAGAAA 530
Db 481 TGTTCAATTAATTTCTCCCAATAAAGCTTTTACAGCCTTTTTCGAAAGAAA 530

RESULT 6
AF055473 528 bp mRNA linear PRI 01-MAY-2000
LOCUS Homo sapiens GAGE-8 mRNA, complete cds.
DEFINITION AF055473
ACCESSION AF055473
VERSION AF055473.1 GI:3511022
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 528)
De Backer, O., Arden, K.C., Boretti, M., Vantomme, V., De Smet, C.,
Czekay, S., Viars, C.S., De Plaen, E., Brasseur, F., Chomez, P., Van den
Eynde, B., Boon, T. and van der Bruggen, P.
Characterization of the GAGE genes that are expressed in various
human cancers and in normal testis
Cancer Res. 59 (13), 3157-3165 (1999)

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
10397259
2 (bases 1 to 528)
De Backer, O.R.Y.
Direct Submission
Submitted (24-MAR-1998) Ludwig Institute for Cancer Research,
Brussels Branch, 74 av. Hippocrate, Brussels B-1200, Belgium
Location/Qualifiers
FEATURES
source 1..528
/organism="Homo sapiens"
/db_xref="taxon:9606"
73..423
/codon_start=1
/product="GAGE-8"
/protein_id="AAC33676.1"
/db_xref="GI:3511023"
/translation="MSWRGRSTYRPRRPRRYVEPPMIGMRPEQFSDVEVPATPEEGE
PATORDPAAOEGEDGASAGQGPKEADSQEQHPQTGCECEDGPDQEMDPNPE
EVKTPPEGEKQSQ"

BASE COUNT 158 a 112 c 151 g 107 t
ORIGIN

Query Match 97.7%; Score 525.4; DB 9; Length 528;
Best Local Similarity 99.8%; Pred. No. 7.3e-117;
Matches 526; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 12 CTGTGAGGAGTCTGTGTGGTTCTCGCGTCCGGACTCTTTTCTCTACTGAGATTCA 71
Db 1 CTGTGAGGAGTCTGTGTGGTTCTCGCGTCCGGACTCTTTTCTCTACTGAGATTCA 60

Qy 72 TCTGTGTGAATATGAGTTGGCGAGGAAGATCGACCTATCGGCCTAGACCAAGACGCTAC 131
Db 61 TCTGTGTGAATATGAGTTGGCGAGGAAGATCGACCTATCGGCCTAGACCAAGACGCTAC 120

Qy 132 GTAGAGCCTCTCTGAATGATTTGGCCCTATGCGGCCCGAGCAGTTTCAGTGCATGAAGTGGAA 191
Db 121 GTAGAGCCTCTCTGAATGATTTGGCCCTATGCGGCCCGAGCAGTTTCAGTGCATGAAGTGGAA 180

Qy 192 CCAGCAACACCTCAAGAAGGGGAACACAGCAACTCAACCTCAGGATCCTGCAGTGCCTCAG 251
Db 181 CCAGCAACACCTCAAGAAGGGGAACACAGCAACTCAACCTCAGGATCCTGCAGTGCCTCAG 240

Qy 252 GAGGGAGAGGATGAGGAGGATCTGAGGTCAAGGCCCGAGCCCTGAGCCTCATATGCCAG 311
Db 241 GAGGGAGAGGATGAGGAGGATCTGAGGTCAAGGCCCGAGCCCTGAGCCTCATATGCCAG 300

Qy 312 GAACAGGTTCAACACACAGACTGGGTGTGAGTGTGAAGATGGTCCCTGATGGCAGGAGATG 371
Db 301 GACAGGGTCAACACACAGACTGGGTGTGAGTGTGAAGATGGTCCCTGATGGCAGGAGATG 360

Qy 372 GACCCGCCCAATCCAGAGAGGTGAACACGCTTGAAGAGGTGAAGAGCAATCACAGTGT 431
Db 361 GACCCGCCCAATCCAGAGAGGTGAACACGCTTGAAGAGGTGAAGAGCAATCACAGTGT 420

Qy 432 TAAAGAAGACACGTTGAAATGATGAGGCTGCTCTATGTTGGAAATTTGTTTCATTAAA 491
Db 421 TAAAGAAGACACGTTGAAATGATGAGGCTGCTCTATGTTGGAAATTTGTTTCATTAAA 480

Qy 492 ATTCTCCCAATAAAGCTTTTACAGCCTTCTGCAAGAGAAAAA 538
Db 481 ATTCTCCCAATAAAGCTTTTACAGCCTTCTGCAAGAGAAAAA 527

RESULT 7

AR028490
LOCUS AR028490 540 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 16 from patent US 5858689.
ACCESSION AR028490
VERSION AR028490.1 GI:5940463
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 540)
AUTHORS van der Bruggen,P., van den Eynde,B., DeBacker,O. and Boon-Falleur,T.
TITLE Isolated peptides derived from the gage tumor rejection antigen precursor and uses thereof
JOURNAL Patent: US 5858689-A 16 12-JAN-1999;
FEATURES Location/Qualifiers
source 1..540
BASE COUNT 159 a 114 c 156 g 111 t
ORIGIN

Query Match 95.3%; Score 512.8; DB 6; Length 540;
Best Local Similarity 98.1%; Pred. No. 8.2e-114;
Matches 530; Conservative 0; Mismatches 7; Indels 3; Gaps 1;

QY 2 CGGCAGGAGCTGTGAGCGAGTGTGTGGTTCTCGCGTCCGGACTCTTTTCCCTCTA 61
Db 1 CGGCAGGAGCTGTGAGCGAGTGTGTGGTTCTCGCGTCCGGACTCTTTTCCCTCTA 60

QY 62 CTGAGATTTCATCTGTGTCGAATATGAGTTGGCGAGGAGATCGACC---TATCGGCCTAG 118
Db 61 CTGAGATTTCATCTGTGTCGAATATGAGTTGGCGAGGAGATCGACCCTATTATTGGCCCTAG 120

QY 119 ACCAAGACGCTACGTAGAGCCTCTCTGAAATGATTGGGCCTATGCGGCCGAGCAGTTTCAG 178
Db 121 ACCAAGCGCTATGTACAGCCTCTCTGAAATGATTGGGCCTATGCGGCCGAGCAGTTTCAG 180

QY 179 TGATGAAGTGAACACAGCAACACCTTGAAGAGGGGAACCAAGCAACTCAACGTCAGGATCC 238
Db 181 TGATGAAGTGAACACAGCAACACCTTGAAGAGGGGAACCAAGCAACTCAACGTCAGGATCC 240

QY 239 TGCAGCTGCTCAGGAGGAGAGGATGAGGAGGATCTGCAAGTCAAGGGCCGAAGCCCTGA 298
Db 241 TGCAGCTGCTCAGGAGGAGAGGATGAGGAGGATCTGCAAGTCAAGGGCCGAAGCCCTGA 300

QY 299 AGCTCATAGCCAGGAACAGGGTCACCCACAGACTGGGTGTGAGTGTGAAGATGGTCCCTGA 358
Db 301 AGCTGATAGCCAGGAACAGGGTCACCCACAGACTGGGTGTGAGTGTGAAGATGGTCCCTGA 360

QY 359 TGGCAGGAGATGGACCGCCAAATCCAGAGAGGTGAACACGGCTTGAAGAGGTGAANA 418
Db 361 TGGCAGGAGATGGACCGCCAAATCCAGAGAGGTGAACACGGCTTGAAGAGGTGAANA 420

QY 419 GCAATCACAGTGTAAAAGAAGACACGTTGAAATGATGACAGGCTGCTCTATGTTGGAAA 478
Db 421 GCAATCACAGTGTAAAAGAAGACACGTTGAAATGATGACAGGCTGCTCTATGTTGGAAA 480

QY 479 TTTGTTCAATTAATAATCTCCCAATAAAGCTTTACAGCCTTCTGCAAAAGAAAAA 538
Db 481 TTTGTTCAATTAATAATCTCCCAATAAAGCTTTACAGCCTTCTGCAAAAGAAAAA 540

RESULT 8
I55853
LOCUS I55853 540 bp DNA linear PAT 07-OCT-1997
DEFINITION Sequence 16 from patent US 5648226.
ACCESSION I55853
VERSION I55853.1 GI:2476647
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.

REFERENCE 1 (bases 1 to 540)
AUTHORS Van den Eynde,B., DeBacker,O. and Boon-Falleur,T.
TITLE Isolated peptides derived from tumor rejection antigens, and their use
JOURNAL Patent: US 5648226-A 16 15-JUL-1997;
FEATURES Location/Qualifiers
source 1..540
BASE COUNT 159 a 114 c 156 g 111 t
ORIGIN

Query Match 95.3%; Score 512.8; DB 6; Length 540;
Best Local Similarity 98.1%; Pred. No. 8.2e-114;
Matches 530; Conservative 0; Mismatches 7; Indels 3; Gaps 1;

QY 2 CGGCAGGAGCTGTGAGCGAGTGTGTGGTTCTCGCGTCCGGACTCTTTTCCCTCTA 61
Db 1 CGGCAGGAGCTGTGAGCGAGTGTGTGGTTCTCGCGTCCGGACTCTTTTCCCTCTA 60

QY 62 CTGAGATTTCATCTGTGTCGAATATGAGTTGGCGAGGAGATCGACC---TATCGGCCTAG 118
Db 61 CTGAGATTTCATCTGTGTCGAATATGAGTTGGCGAGGAGATCGACCCTATTATTGGCCCTAG 120

QY 119 ACCAAGACGCTACGTAGAGCCTCTCTGAAATGATTGGGCCTATGCGGCCGAGCAGTTTCAG 178
Db 121 ACCAAGGGCTATGTACAGCCTCTCTGAAATGATTGGGCCTATGCGGCCGAGCAGTTTCAG 180

QY 179 TGATGAAGTGAACACAGCAACACCTTGAAGAGGGGAACCAAGCAACTCAACGTCAGGATCC 238
Db 181 TGATGAAGTGAACACAGCAACACCTTGAAGAGGGGAACCAAGCAACTCAACGTCAGGATCC 240

QY 239 TGCAGCTGCTCAGGAGGAGAGGATGAGGAGGATCTGCAAGTCAAGGGCCGAAGCCCTGA 298
Db 241 TGCAGCTGCTCAGGAGGAGAGGATGAGGAGGATCTGCAAGTCAAGGGCCGAAGCCCTGA 300

QY 299 AGCTCATAGCCAGGAACAGGGTCACCCACAGACTGGGTGTGAGTGTGAAGATGGTCCCTGA 358
Db 301 AGCTGATAGCCAGGAACAGGGTCACCCACAGACTGGGTGTGAGTGTGAAGATGGTCCCTGA 360

QY 359 TGGCAGGAGATGGACCGCCAAATCCAGAGAGGTGAACACGGCTTGAAGAGGTGAANA 418
Db 361 TGGCAGGAGATGGACCGCCAAATCCAGAGAGGTGAACACGGCTTGAAGAGGTGAANA 420

QY 419 GCAATCACAGTGTAAAAGAAGACACGTTGAAATGATGACAGGCTGCTCTATGTTGGAAA 478
Db 421 GCAATCACAGTGTAAAAGAAGACACGTTGAAATGATGACAGGCTGCTCTATGTTGGAAA 480

QY 479 TTTGTTCAATTAATAATCTCCCAATAAAGCTTTACAGCCTTCTGCAAAAGAAAAA 538
Db 481 TTTGTTCAATTAATAATCTCCCAATAAAGCTTTACAGCCTTCTGCAAAAGAAAAA 540

RESULT 9
AR028492
LOCUS AR028492 539 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 18 from patent US 5858689.
ACCESSION AR028492
VERSION AR028492.1 GI:5940465
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 539)
AUTHORS van der Bruggen,P., van den Eynde,B., DeBacker,O. and Boon-Falleur,T.
TITLE Isolated peptides derived from the gage tumor rejection antigen precursor and uses thereof
JOURNAL Patent: US 5858689-A 18 12-JAN-1999;
FEATURES Location/Qualifiers
source 1..539
BASE COUNT 158 a 113 c 157 g 111 t
ORIGIN

ORIGIN

Query Match 94.8%; Score 510.2; DB 6; Length 539;
Best Local Similarity 98.0%; Pred. No. 3.5e-113;
Matches 528; Conservative 0; Mismatches 8; Indels 3; Gaps 1;

QY 3 GCCAGGAGCTGTGAGGCAGTCTGTGTGTTCTGCGCGCGGACCTCTTTTCTCTAC 62
DB 1 GCCAGGAGCTGTGAGGCAGTCTGTGTGTTCTGCGCGCGGACCTCTTTTCTCTAC 60

QY 63 TGAGATTTCATCTGTGTGAAATATAGTTGGCGAGGAAGATCGACC---TATCGGCCTAGA 119
DB 61 TGAGATTTCATCTGTGTGAAATATAGTTGGCGAGGAAGATCGACCTATTATTGGCCCTAGA 120

QY 120 CCAAGAGCCTACGTAGAGCCTCTGAAATGATTGGGCCCTATGCGGCCGAGCAGTTCACT 179
DB 121 CCAAGGCGCTATGTACAGCCCTCTGAAAGTATTGGGCCCTATGCGGCCGAGCAGTTCACT 180

QY 180 GATGAAGTGGAAACCAACACCTCTGAAAGAGGGGAACACAGCAACTCAACGTCAGGATCCT 239
DB 181 GATGAAGTGGAAACCAACACCTCTGAAAGAGGGGAACACAGCAACTCAACGTCAGGATCCT 240

QY 240 GCAGCTCTCAGGAGGAGAGATGAGGAGCATCTGCAGGTCGAAGGCCGAAGCCTGAA 299
DB 241 GCAGCTCTCAGGAGGAGAGATGAGGAGCATCTGCAGGTCGAAGGCCGAAGCCTGAA 300

QY 300 GCTCATAGCCAGGAACAGGGTCACCCACAGACTGGGTGTGAGTGTGAAGATGTCCTGAT 359
DB 301 GCTGATAGCCAGGAACAGGGTCACCCACAGACTGGGTGTGAGTGTGAAGATGTCCTGAT 360

QY 360 GGCAGGAGATGGAACCGCCAAATCCAGAGAGGTGAAACCGCTCGAAGAGGTGAAAG 419
DB 361 GGCAGGAGGTGGACCGCCAAATCCAGAGAGGTGAAACCGCTCGAAGAGGTGAAAG 420

QY 420 CAATCACAGTGTAAAAGAACACAGTTCGAAATGATGCAGGCTGCTCTATGTTGGAAT 479
DB 421 CAATCACAGTGTAAAAGAACACAGTTCGAAATGATGCAGGCTGCTCTATGTTGGAAT 480

QY 480 TTGTTTCATTAAATTCCTCCCAATAAAGCTTTACAGCCTTCTGCAAGAAAAA 538
DB 481 TTGTTTCATTAAATTCCTCCCAATAAAGCTTTACAGCCTTCTGCAAGAAAAA 539

RESULT 10
LOCUS I55855 539 bp DNA linear PAT 07-OCT-1997
DEFINITION Sequence 18 from patent US 5648226.
ACCESSION I55855
VERSION I55855.1 GI:2476649
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 539)
AUTHORS Van den Eynde,B., DeBacker,O. and Boon-Falleur,T.
TITLE Isolated peptides derived from tumor rejection antigens, and their use
JOURNAL Patent: US 5648226-A 18 15-JUL-1997;
FEATURES Location/Qualifiers
source 1..539
BASE COUNT 158 a 113 c 157 g 111 t
ORIGIN

Query Match 94.8%; Score 510.2; DB 6; Length 539;
Best Local Similarity 98.0%; Pred. No. 3.5e-113;
Matches 528; Conservative 0; Mismatches 8; Indels 3; Gaps 1;

QY 3 GCCAGGAGCTGTGAGGCAGTCTGTGTGTTCTGCGCGCGGACCTCTTTTCTCTAC 62
DB 1 GCCAGGAGCTGTGAGGCAGTCTGTGTGTTCTGCGCGCGGACCTCTTTTCTCTAC 60

QY 63 TGAGATTTCATCTGTGTGAAATATAGTTGGCGAGGAAGATCGACC---TATCGGCCTAGA 119
DB 61 TGAGATTTCATCTGTGTGAAATATAGTTGGCGAGGAAGATCGACCTATTATTGGCCCTAGA 120

QY 120 CCAAGAGCCTACGTAGAGCCTCTGAAATGATTGGGCCCTATGCGGCCGAGCAGTTCACT 179
DB 121 CCAAGGCGCTATGTACAGCCCTCTGAAAGTATTGGGCCCTATGCGGCCGAGCAGTTCACT 180

QY 180 GATGAAGTGGAAACCAACACCTCTGAAAGAGGGGAACACAGCAACTCAACGTCAGGATCCT 239
DB 181 GATGAAGTGGAAACCAACACCTCTGAAAGAGGGGAACACAGCAACTCAACGTCAGGATCCT 240

QY 240 GCAGCTCTCAGGAGGAGAGATGAGGAGCATCTGCAGGTCGAAGGCCGAAGCCTGAA 299
DB 241 GCAGCTCTCAGGAGGAGAGATGAGGAGCATCTGCAGGTCGAAGGCCGAAGCCTGAA 300

QY 300 GCTCATAGCCAGGAACAGGGTCACCCACAGACTGGGTGTGAGTGTGAAGATGTCCTGAT 359
DB 301 GCTGATAGCCAGGAACAGGGTCACCCACAGACTGGGTGTGAGTGTGAAGATGTCCTGAT 360

QY 360 GGCAGGAGATGGAACCGCCAAATCCAGAGAGGTGAAACCGCTCGAAGAGGTGAAAG 419
DB 361 GGCAGGAGGTGGACCGCCAAATCCAGAGAGGTGAAACCGCTCGAAGAGGTGAAAG 420

QY 420 CAATCACAGTGTAAAAGAACACAGTTCGAAATGATGCAGGCTGCTCTATGTTGGAAT 479
DB 421 CAATCACAGTGTAAAAGAACACAGTTCGAAATGATGCAGGCTGCTCTATGTTGGAAT 480

QY 480 TTGTTTCATTAAATTCCTCCCAATAAAGCTTTACAGCCTTCTGCAAGAAAAA 538
DB 481 TTGTTTCATTAAATTCCTCCCAATAAAGCTTTACAGCCTTCTGCAAGAAAAA 539

RESULT 11
LOCUS AR028491 532 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 17 from patent US 5858689.
ACCESSION AR028491
VERSION AR028491.1 GI:5940464
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 532)
AUTHORS van der Bruggen,P., van den Eynde,B., DeBacker,O. and Boon-Falleur,T.
TITLE Isolated peptides derived from the gage tumor rejection antigen precursor and uses thereof
JOURNAL Patent: US 5858689-A 17 12-JAN-1999;
FEATURES Location/Qualifiers
source 1..532
BASE COUNT 156 a 111 c 154 g 111 t
ORIGIN

Query Match 93.8%; Score 504.8; DB 6; Length 532;
Best Local Similarity 98.1%; Pred. No. 7.1e-112;
Matches 522; Conservative 0; Mismatches 7; Indels 3; Gaps 1;

QY 10 AGCTGTGAGGCAGTGTGTGTTCTGCGCGTCCGGACTCTTTTCTCTACTGAGATT 69
DB 1 AGCTGTGAGGCAGTGTGTGTTCTGCGCGTCCGGACTCTTTTCTCTACTGAGATT 60

QY 70 CATCTCTGTGAAATATAGTTGGCGAGGAAGATCGACC---TATCGGCCTAGACCAAGAC 126
DB 61 CATCTCTGTGAAATATAGTTGGCGAGGAAGATCGACCATTATTATGGCCCTAGACCAAGGC 120

QY 127 GCTACGTAGAGCCTCTGAAATGATTGGCCCTATGCGGCCCGCAGCAGTTTCAGTGAAG 186
DB 121 GCTATGTACAGCCTCTCTGAAAGTATTGGCCCTATGCGGCCCGCAGCAGTTTCAGTGAAG 180

QY 63 TGAGATTTCATCTGTGTGAAATATAGTTGGCGAGGAAGATCGACC---TATCGGCCTAGA 119
DB 61 TGAGATTTCATCTGTGTGAAATATAGTTGGCGAGGAAGATCGACCTATTATTGGCCCTAGA 120

QY 120 CCAAGAGCCTACGTAGAGCCTCTGAAATGATTGGGCCCTATGCGGCCGAGCAGTTCACT 179
DB 121 CCAAGGCGCTATGTACAGCCCTCTGAAAGTATTGGGCCCTATGCGGCCGAGCAGTTCACT 180

QY 180 GATGAAGTGGAAACCAACACCTCTGAAAGAGGGGAACACAGCAACTCAACGTCAGGATCCT 239
DB 181 GATGAAGTGGAAACCAACACCTCTGAAAGAGGGGAACACAGCAACTCAACGTCAGGATCCT 240

QY 240 GCAGCTCTCAGGAGGAGAGATGAGGAGCATCTGCAGGTCGAAGGCCGAAGCCTGAA 299
DB 241 GCAGCTCTCAGGAGGAGAGATGAGGAGCATCTGCAGGTCGAAGGCCGAAGCCTGAA 300

QY 300 GCTCATAGCCAGGAACAGGGTCACCCACAGACTGGGTGTGAGTGTGAAGATGTCCTGAT 359
DB 301 GCTGATAGCCAGGAACAGGGTCACCCACAGACTGGGTGTGAGTGTGAAGATGTCCTGAT 360

QY 360 GGCAGGAGATGGAACCGCCAAATCCAGAGAGGTGAAACCGCTCGAAGAGGTGAAAG 419
DB 361 GGCAGGAGGTGGACCGCCAAATCCAGAGAGGTGAAACCGCTCGAAGAGGTGAAAG 420

QY 420 CAATCACAGTGTAAAAGAACACAGTTCGAAATGATGCAGGCTGCTCTATGTTGGAAT 479
DB 421 CAATCACAGTGTAAAAGAACACAGTTCGAAATGATGCAGGCTGCTCTATGTTGGAAT 480

QY 480 TTGTTTCATTAAATTCCTCCCAATAAAGCTTTACAGCCTTCTGCAAGAAAAA 538
DB 481 TTGTTTCATTAAATTCCTCCCAATAAAGCTTTACAGCCTTCTGCAAGAAAAA 539

RESULT 11
LOCUS AR028491 532 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 17 from patent US 5858689.
ACCESSION AR028491
VERSION AR028491.1 GI:5940464
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 532)
AUTHORS van der Bruggen,P., van den Eynde,B., DeBacker,O. and Boon-Falleur,T.
TITLE Isolated peptides derived from the gage tumor rejection antigen precursor and uses thereof
JOURNAL Patent: US 5858689-A 17 12-JAN-1999;
FEATURES Location/Qualifiers
source 1..532
BASE COUNT 156 a 111 c 154 g 111 t
ORIGIN

Query Match 93.8%; Score 504.8; DB 6; Length 532;
Best Local Similarity 98.1%; Pred. No. 7.1e-112;
Matches 522; Conservative 0; Mismatches 7; Indels 3; Gaps 1;

QY 10 AGCTGTGAGGCAGTGTGTGTTCTGCGCGTCCGGACTCTTTTCTCTACTGAGATT 69
DB 1 AGCTGTGAGGCAGTGTGTGTTCTGCGCGTCCGGACTCTTTTCTCTACTGAGATT 60

QY 70 CATCTCTGTGAAATATAGTTGGCGAGGAAGATCGACC---TATCGGCCTAGACCAAGAC 126
DB 61 CATCTCTGTGAAATATAGTTGGCGAGGAAGATCGACCATTATTATGGCCCTAGACCAAGGC 120

QY 127 GCTACGTAGAGCCTCTGAAATGATTGGCCCTATGCGGCCCGCAGCAGTTTCAGTGAAG 186
DB 121 GCTATGTACAGCCTCTCTGAAAGTATTGGCCCTATGCGGCCCGCAGCAGTTTCAGTGAAG 180

QY 167 TCGAACCAACACCTGAAGAAGGGGAACAGCAACTCAACGTGAGGATCCTGACGCTG 246
|||||
Db 181 TGAACCAACCAACCTGAAGAAGGGGAACAGCAACTCAACGTGAGGATCCTGACGCTG 240
QY 247 CTCAGAGGAGAGGATGAGGGAGCATCTGCAGGTCAAGGGCCGAAGCCTGAAGCTCATA 306
|||||
Db 241 CTCAGAGGAGAGGATGAGGGAGCATCTGCAGGTCAAGGGCCGAAGCCTGAAGCTGATA 300
QY 307 GCCAGAACAGGTCACCCACAGACTGGGTGTGAGTGTGAAGATGCTCCTGATGGGCAGG 366
|||||
Db 301 GCCAGAACAGGTCACCCACAGACTGGGTGTGAGTGTGAAGATGCTCCTGATGGGCAGG 360
QY 367 AGATGACCCGCCAAATCCAGAGGAGGTGAAACCCCTGACAGAGTGAAGCAATCAC 426
|||||
Db 361 AGATGACCCGCCAAATCCAGAGGAGGTGAAACCCCTGACAGAGTGAAGCAATCAC 420
QY 427 AGTGTAAAGAACAGACACGTTGAAATGATGACAGCTGCTCTATGTTGAAATTTGTTCA 486
|||||
Db 421 AGTGTAAAGAACAGACACGTTGAAATGATGACAGCTGCTCTATGTTGAAATTTGTTCA 480
QY 487 TTAATAATCTCCCAATAAAGCTTTACAGCCTTCTGCAAGAGAAAAA 538
|||||
Db 481 TTAATAATCTCCCAATAAAGCTTTACAGCCTTCTGCAAGAGAAAAA 532

RESULT 12
LOCUS I55854 532 bp DNA linear PAT 07-OCT-1997
DEFINITION Sequence 17 from patent US 5648226.
ACCESSION I55854
VERSION I55854.1 GI:2476648
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 532)
AUTHORS Van den Eynde,B., DeBacker,O. and Boon-Falleur,T.
TITLE Isolated peptides derived from tumor rejection antigens, and their use
JOURNAL Patent: US 5648226-A 17 15-JUL-1997;
FEATURES
source Location/Qualifiers
1..532
BASE COUNT 156 a 111 c 154 g 111 t
ORIGIN

Query Match 93.8%; Score 504.8; DB 6; Length 532;
Best Local Similarity 98.1%; Pred. No. 7.1e-112;
Matches 522; Conservative 0; Mismatches 7; Indels 3; Gaps 1;

QY 10 AGCTGTGAGGCAGTCTGTGTGTTCTGCGTCCGCGACTCTTTTCTCTACTGAGATT 69
|||||
Db 1 AGCTGTGAGGCAGTCTGTGTGTTCTGCGTCCGCGACTCTTTTCTCTACTGAGATT 60
QY 70 CATCTGTGAAATATGAGTTGGCGAGGAAGATCGACCC---TATCGGCTAGACCAAGAC 126
|||||
Db 61 CATCTGTGAAATATGAGTTGGCGAGGAAGATCGACCCATATTATGGCCTAGACCAAGC 120
QY 127 GCTACGTAGAGCCTCTGAAATGATTGGGCCTATATGGCCCGGAGCAGTTGATGAAG 186
|||||
Db 121 GCTATGTACAGCCTCTGAAATGATTGGGCCTATATGGCCCGGAGCAGTTGATGAAG 180
QY 187 TGGAACCAAGCAACCTGAAAGAGGGGAACCACTCAACGTGAGGATCCTGACGCTG 246
|||||
Db 181 TGGAACCAAGCAACCTGAAAGAGGGGAACCACTCAACGTGAGGATCCTGACGCTG 240
QY 247 CTCAGGAGGAGGATGAGGAGCATCTGACAGTCAAGGGCCGAAGCCTGAAGCTCATA 306
|||||
Db 241 CTCAGGAGGAGGATGAGGAGCATCTGACAGTCAAGGGCCGAAGCCTGAAGCTGATA 300
QY 307 GCCAGGAACAGGTCACCCACAGACTGGGTGTGAGTGTGAAGATGTTGCTGATGGGCAGG 366
|||||

Db 301 GCCAGGAACAGGGTCAACCCACAGACTGGGTGTGAGTGTGAAGATGTTCTGTGGCAGG 360
QY 367 AGATGACCCGCCAAATCCAGAGGAGGTGAAACCCCTCAAGCAAGTGAAGCAATCAC 426
|||||
Db 361 AGATGACCCGCCAAATCCAGAGGAGGTGAAACCCCTCAAGCAAGTGAAGCAATCAC 420
QY 427 AGTGTAAAGAACAGACACGTTGAAATGATGACAGGCTGCTCTATGTTGAAATTTGTTCA 486
|||||
Db 421 AGTGTAAAGAACAGGACGTTGAAATGATGACAGGCTGCTCTATGTTGAAATTTGTTCA 480
QY 487 TTAATAATCTCCCAATAAAGCTTTACAGCCTTCTGCAAGAGAAAAA 538
|||||
Db 481 TTAATAATCTCCCAATAAAGCTTTACAGCCTTCTGCAAGAGAAAAA 532

RESULT 13
LOCUS AX334151 528 bp DNA linear PAT 09-JAN-2002
DEFINITION Sequence 4660 from Patent WO0194629.
ACCESSION AX334151
VERSION AX334151.1 GI:18124870
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (sites)
AUTHORS Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endress,G.,
Horrigan,S., Soppet,D.R. and Weaver,Z.
TITLE Cancer gene determination and therapeutic screening using signature gene sets
JOURNAL Patent: WO 0194629-A 4660 13-DEC-2001;
FEATURES
source Location/Qualifiers
1..528
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 147 a 114 c 156 g 111 t
ORIGIN

Query Match 93.2%; Score 501.4; DB 6; Length 528;
Best Local Similarity 98.3%; Pred. No. 4.7e-111;
Matches 518; Conservative 0; Mismatches 6; Indels 3; Gaps 1;

QY 2 CGCCAGGAGCTGTGAGGAGTCTGTGTGTTCTGCGTCCGCGACTCTTTTCTCTCTA 61
|||||
Db 1 CGCCAGGAGCTGTGAGGAGTCTGTGTGTTCTGCGTCCGCGACTCTTTTCTCTCTA 60
QY 62 CTGAGATTCTGTGTGAAATATGAGTTGGCGAGGAAGATCGACCC---TATCGGCTAG 118
|||||
Db 61 CTGAGATTCTGTGTGAAATATGAGTTGGCGAGGAAGATCGACCTATTATTGGCCTAG 120
QY 119 ACCAAGAGCTAGTAGAGCCTCTGAAATGATTGGGCCTATGCGGCCCGGAGCAGTTGAG 178
|||||
Db 121 ACCAAGGCCTATGTACAGCCTCTGAAATGATTGGGCCTATGCGGCCCGGAGCAGTTGAG 180
QY 179 TGATGAAGTGGAAACCAACACCTGAAAGAGGGGAACAGCAACTCAACGTGAGGATCC 238
|||||
Db 181 TGATGAAGTGGAAACCAACACCTGAAAGAGGGGAACCACTCAACGTGAGGATCC 240
QY 239 TGCAGCTGCTCAGGAGGAGAGATGAGGGAGCATCTGACAGTCAAGGGCCGAAGCCTGA 298
|||||
Db 241 TGCAGCTGCTCAGGAGGAGAGATGAGGGAGCATCTGACAGTCAAGGGCCGAAGCCTGA 300
QY 299 AGCTCATAGCCAGGAACAGGGTCACCCACAGACTGGGTGTGAGTGTGAAGATGGTCTCTCA 358
|||||
Db 301 AGCTGATAGCCAGGAACAGGGTCACCCACAGACTGGGTGTGAGTGTGAAGATGGTCTCTCA 360
QY 359 TGGCAGGAGATGGACCCGCCAAATCCAGAGGAGGTGAAACCGCTGAAGAGGTGAAAA 418
|||||
Db 361 TGGCAGGAGATGGACCCGCCAAATCCAGAGGAGGTGAAACCGCTGAAGAGGTGAAAA 420
|||||

QY 419 GCAATCACAGTGTAAAGAGACACAGTGAATGATGAGGCTGCTCTCTATTGGAAA 478
|||||
Db 421 GCAATCACAGTGTAAAGAGAGCAGCTGGAATGATGAGGCTGCTCTATTGGAAA 480
|||||
QY 479 TTTGTTCAATTAATCTCCCAATAAAGCTTTACAGCCTTCGCAA 525
|||||
Db 481 TTTGTTCAATTAATCTCCCAATAAAGCTTTACAGCCTTCGCAA 527
|||||
RESULT 14
HSU19145 528 bp mRNA linear PRI 04-DEC-1995
LOCUS Human GAGE-4 protein mRNA, complete cds.
ACCESSION U19145
VERSION U19145.1 GI:914904
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 528)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Van den Eynde,B., Peeters,O., De Backer,O., Gaugler,B., Lucas,S.
and Boon,T.
TITLE A new family of genes coding for an antigen recognized by
autologous cytolytic T lymphocytes on a human melanoma
J. Exp. Med. 182 (3), 689-698 (1995)
JOURNAL 95378788
REFERENCE 2 (bases 1 to 528)
AUTHORS Van Den Eynde,B.J.
TITLE Direct Submission
JOURNAL Submitted (28-DEC-1994) BENOIT J VAN DEN EYNDE, Ludwig Institute
for Cancer Research, 74 Avenue Hippocrate, Brussels B-1200, Belgium
FEATURES
source
1. .528
/organism="Homo sapiens"
/db_xref="taxon:9606"
/sex="female"
/cell_line="M2-MEL.43"
/tissue_type="melanoma"
/dev_stage="adult"
83. .436
/codon_start=1
/product="GAGE-4 protein"
/protein_id="AAA82747.1"
/db_xref="GI:914905"
/translation="MSWRGRSTYYWPRPRRYVQPPMIGPMRPEQSFDEVEPATPEEG
EPATQRDPAAOGEDEGASAGQPKPEADSQEKGHPQTGCECEDGPDQEMDPNP
EEVKTPPEGEKQSQ"
BASE COUNT 147 a 114 c 156 g 111 t
ORIGIN
Query Match 93.2%; Score 501.4; DB 9; Length 528;
Best Local Similarity 98.3%; Pred. No. 4.7e-111;
Matches 518; Conservative 0; Mismatches 6; Indels 3; Gaps 1;
QY 2 CGCCAGGAGCTGTGAGGACGTGTGTGTTCTCTCGCGTCGGGACTCTTTTCTCTA 61
|||||
Db 1 CGCCAGGAGCTGTGAGGACGTGTGTGTTCTCTCGCGTCGGGACTCTTTTCTCTA 60
|||||
QY 62 CTGAGATTCACTGTGAAATATGAGTTGGCGAGAGATCGACC---TATCGGCCTAG 118
|||||
Db 61 CTGAGATTCACTGTGAAATATGAGTTGGCGAGAGATCGACCATTATTGGCCTAG 120
|||||
QY 119 ACCAAGCGCTACTAGACCTCTCTGAATGATTGGCTATCGCGCCGAGCAGTTTCAG 178
|||||
Db 121 ACCAAGCGCTATGATACCTCTCTGAATGATTGGCTATGCGGCCCGAGCAGTTTCAG 180
|||||
QY 179 TGATGAAGTGAACACGACCACTGGAAGAGGGGAACACGAACTCAACGTCAGGATCC 238
|||||
Db 181 TGATGAAGTGAACACCACTGGAAGAGGGGAACACGAACTCAACGTCAGGATCC 240
|||||
QY 239 TGAGAGTGTCTAGAGGAGAGGATGAGGGAGCATCTGCGAGTCAAGGCCGAGCCTGA 298
|||||

Db 241 TCAGAGTGTCTCAGAGGGAGAGGATGAGGGAGCATCTGCAGTCCAGGCCGAGGCTGA 300
|||||
QY 299 AGCTCATAGCCAGGAGGAGGTCACCCACACACACAGTGGGTGTGAGTGTGAAGATGCTCTGA 358
|||||
Db 301 AGCTCATAGCCAGGAGGAGGTCACCCACACACAGTGGGTGTGAGTGTGAAGATGCTCTGA 360
|||||
QY 359 TGGGAGGAGATGGACCCGCCAAATCCAGAGGAGGTGAAAACGCCCTGAAAGAGGTGAAAA 418
|||||
Db 361 TGGGAGGAGATGGACCCGCCAAATCCAGAGGAGGTGAAAACGCCCTGAAAGAGGTGAAAA 420
|||||
QY 419 GCAATCACAGTGTAAAGAGACACAGTGTGAATGATGAGGCTGCTCTATTTGGAAA 478
|||||
Db 421 GCAATCACAGTGTAAAGAGAGCAGCTTGAATGATGAGGCTGCTCTATTTGGAAA 480
|||||
QY 479 TTTGTTCAATTAATCTCTCCCAATAAAGCTTTACAGCCTTCGCAA 525
|||||
Db 481 TTTGTTCAATTAATCTCTCCCAATAAAGCTTTACAGCCTTCGCAA 527
|||||
RESULT 15
HSU19147 527 bp mRNA linear PRI 04-DEC-1995
LOCUS Human GAGE-6 protein mRNA, complete cds.
ACCESSION U19147
VERSION U19147.1 GI:914908
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 527)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Van den Eynde,B., Peeters,O., De Backer,O., Gaugler,B., Lucas,S.
and Boon,T.
TITLE A new family of genes coding for an antigen recognized by
autologous cytolytic T lymphocytes on a human melanoma
J. Exp. Med. 182 (3), 689-698 (1995)
JOURNAL 95378788
REFERENCE 2 (bases 1 to 527)
AUTHORS Van Den Eynde,B.J.
TITLE Direct Submission
JOURNAL Submitted (28-DEC-1994) BENOIT J VAN DEN EYNDE, Ludwig Institute
for Cancer Research, 74 Avenue Hippocrate, Brussels B-1200, Belgium
FEATURES
source
1. .527
/organism="Homo sapiens"
/db_xref="taxon:9606"
/sex="female"
/cell_line="M2-MEL.43"
/tissue_type="melanoma"
/dev_stage="adult"
82. .435
/codon_start=1
/product="GAGE-6 protein"
/protein_id="AAA82749.1"
/db_xref="GI:914909"
/translation="MSWRGRSTYYWPRPRRYVOPPEVIGPMRPEQSFDEVEPATPEEG
EPATQRDPAAOGEDEGASAGQPKPEADSQEKGHPQTGCECEDGPDQEVDPNP
EEVKTPPEGEKQSQ"
BASE COUNT 146 a 113 c 157 g 111 t
ORIGIN
Query Match 92.7%; Score 498.8; DB 9; Length 527;
Best Local Similarity 98.1%; Pred. No. 2e-110;
Matches 516; Conservative 0; Mismatches 7; Indels 3; Gaps 1;
QY 3 GCAGGAGCTGTGAGGAGCTGTGTGTTCTCTCGCGTCGGGACTCTTTTCTCTAC 62
|||||
Db 1 GCAGGAGCTGTGAGGAGCTGTGTGTTCTCTCGCGTCGGGACTCTTTTCTCTAC 60
|||||
QY 63 TGAGATTCACTGTGAAATATGAGTTGGCGAGGAGATCGACC---TATCGGCCTAGA 119
|||||

```

61  TGAGATTTCATCTGCTGAAATATGAGTTGGCGAGGAAGATCGACCTATTATTGGCCCTAGA 120
Qy 120  CCAAGACGCTACGTAGAGCCTCCTGAAATGATTTGGCCCTATGCGGCCCGGAGCAGTTCACT 179
Db 121  CCAAGGCGCTATGTACAGCCTCCTGAAGTGTGTTGGCCCTATGCGGCCCGGAGCAGTTCACT 180
Qy 180  GATGAAGTGGAAACCAAGCAACACTGAAGAAGGGGAACCAAGCAACTCAACGTCAGGATCCT 239
Db 181  GATGAAGTGGAAACCAAGCAACACTGAAGAAGGGGAACCAAGCAACTCAACGTCAGGATCCT 240
Qy 240  GCAGCTGCTCAGGAGGAGAGGATGAGGAGGATCTGCAGGTCAGGGCCGAGCCCTGAA 299
Db 241  GCAGCTGCTCAGGAGGAGAGGATGAGGAGGATCTGCAGGTCAGGGCCGAGCCCTGAA 300
Qy 300  GCTATAGCCAGGAACAGGGTCACCCACAGACTGGGTGTGAGTCTGAAGATGGTCTTGAT 359
Db 301  GCTATAGCCAGGAACAGGGTCACCCACAGACTGGGTGTGAGTCTGAAGATGGTCTTGAT 360
Qy 360  GGGCAGGAGATGGACCCGCCAAATCCAGAGGAGGTGAAAACGGCTGAAGAAGGTGAAAAG 419
Db 361  GGGCAGGAGGTGGACCCGCCAAATCCAGAGGAGGTGAAAACGGCTGAAGAAGGTGAAAAG 420
Qy 420  CAATCACAGTGTAAAGAAGACACGTTGAAATGATGCAGGCTGCTCCTATGTTGGAAAT 479
Db 421  CAATCACAGTGTAAAGAAGACACGTTGAAATGATGCAGGCTGCTCCTATGTTGGAAAT 480
Qy 480  TTGTTCAATTAATAATCTCCCAATAAAGCTTTACAGCCCTTCTGCAAA 525
Db 481  TTGTTCAATTAATAATCTCCCAATAAAGCTTTACAGCCCTTCTGCAAA 526
```

Search completed: July 1, 2002, 07:13:58
Job time: 3733 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 1, 2002, 06:11:45 ; Search time 1625.58 Seconds
(without alignments)
4466.936 Million cell updates/sec

Title: US-09-782-745-14
Perfect score: 538
Sequence: 1 ACCCGAGGAGCTGTGAGGC.....CTGCAAGAAAAA 538

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_estl:*
10: gb_est2:*
11: gb_hic:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	524.6	97.5	789	10	BG700165
2	492.6	91.6	813	10	BG773070
3	492	91.4	527	10	B1826605
4	492	91.4	851	10	BG250953
5	490.6	91.2	572	10	BG120838
6	476.8	88.6	509	10	B1868671
7	463.4	86.1	623	10	BM172302
8	452.4	84.1	489	9	AA447559
9	431.4	80.2	457	9	AA510753
10	424.6	78.9	447	9	AI381509
11	424.4	78.9	464	9	AA738037
12	424	78.8	581	10	BG700088
13	413.8	76.9	418	9	AW016546
14	413.4	76.8	480	10	BG120336
15	412.8	76.7	426	9	AA868226
16	410.8	76.4	450	9	AA760996
17	401.2	74.6	455	9	AI187350

18	370.4	68.8	795	10	B1560935	BI560935	6032543321
c	19	365.4	67.9	412	9	AW102587	xd67907.x
	20	360.4	67.0	445	11	BC005363	BC005363 Homo sapi
	21	359.4	66.8	419	9	AA448542	zw8le11.r
c	22	346.6	64.4	839	10	BG527731	602556670
	23	329.8	61.3	384	9	AA913206	Om53f03.s
	24	318.8	59.3	397	10	BG206349	RST25797
25	309.8	57.6	1069	10	BE777698	BE777698	
c	26	296.4	55.1	398	9	AA918604	601348269
	27	250.4	46.5	256	10	BG181480	0153f02.s
	28	233.4	43.4	258	10	BG186708	RST328.AC
c	29	228.2	42.4	253	10	BG184057	RST5683
	30	226.6	42.1	261	10	BG199060	RST2973.A
	31	224.8	41.8	245	10	BG212621	BG199060 RST18338
c	32	219.8	40.9	292	10	BG212622	RST32216
	33	214.8	39.9	245	10	BG208433	RST27934
	34	213.6	39.7	849	10	BG436305	BG20508742
c	35	213.4	39.7	524	11	AF318372	Homo sapi
	36	207.6	38.6	507	9	AA972716	op90e12.s
	37	207	38.5	256	10	BG216461	BG216461 RST36040
c	38	204.8	38.1	224	9	AA738394	Yw73e12.r1
	39	202.4	37.6	382	10	BF669799	IL3-ET011
	40	200.6	37.3	499	10	BG354572	IL3-ET011
c	41	193.8	36.0	197	9	AI988311	wu01f07.x
	42	192.8	35.8	226	10	BG220441	RST40225
	43	186.8	34.7	503	9	AI742551	wg55c11.x
c	44	185.8	34.5	756	10	BE903565	601677770
	45	185.8	34.5	756	10	BE903565	601677770

ALIGNMENTS

RESULT 1
BG700165 569 bp mRNA linear EST 07-MAY-2001
LOCUS 602679622F1 NIH_MGC_95 Homo sapiens CDNA clone IMAGE:4812437 5',
DEFINITION mRNA sequence.

ACCESSION BG700165
VERSION BG700165.1 GI:13969233
KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 569)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov

Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM10704 row: 1 column: 06
High quality sequence stop: 566.

FEATURES
Location/Qualifiers
1..569
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4812437"
/clone_lib="NIH_MGC_95"
/tissue_type="hippocampus"
/lab_host="DH10B"
/note="Organ: brain; Vector: pBluescriptR (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',
size-selected for average insert size 2.5 kb and

normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIMH/NHGRI, National Institutes of Health). Note: this is a NIH_MGC Library."

BASE COUNT	166 a	125 c	165 g	113 t
ORIGIN				
Query Match	97.5%	Score 524.6	DB 10	Length 569
Best Local Similarity	99.2%	Pred. No. 1e-111		
Matches 527	Conservative 0	Mismatches 4	Indels 0	Gaps 0
Qy	8	GGAGCTGTGAGGAGTGCTGTGGTTCCTGCGCTCGGACTCTTTTCTCTACTGAGA	67	
Db	25	GGAGCTGTGAGGAGTGCTGTGGTTCCTGCGCTCGGACTCTTTTCTCTACTGAGA	84	
Qy	68	TTTATCTGTGTGAATATGATGTTGGCGAGGAAGATCGACCTATCGGCCTAGACCAAGAGC	127	
Db	85	TTTATCTGTGTGAATATGATGTTGGCGAGGAAGATCGACCTATCGGCCTAGACCAAGAGC	144	
Qy	128	CTACGTAGAGCTCTCTGAAATGATTTGGGCTATCGGCCGAGCAGTTCAGTGATCAAGT	187	
Db	145	CTACGTAGAGCTCTCTGAAATGATTTGGGCTATCGGCCGAGCAGTTCAGTGATCAAGT	204	
Qy	188	GGAAACGAGCAACACCTTGAAGAGGGGAACGAGCAACTCAAGCTCAGGATCTCTGAGCTGC	247	
Db	205	GGAAACGAGCAACACCTTGAAGAGGGGAACGAGCAACTCAAGCTCAGGATCTCTGAGCTGC	264	
Qy	248	TCAGAGGAGAGGATGAGGAGCATCTGCAGGTCAGGCGCGAAGCCTGAAGCTCATAG	307	
Db	265	TCAGAGGAGAGGATGAGGAGCATCTGCAGGTCAGGCGCGAAGCCTGAAGCTCATAG	324	
Qy	308	CCAGGAACAGGGTCACCCACAGACTGGGTGTGAGTGTGAAGATGCTCTGATGGGCAGGA	367	
Db	325	CCAGGAACAGGGTCACCCACAGACTGGGTGTGAGTGTGAAGATGCTCTGATGGGCAGGA	384	
Qy	368	GATGACCGCCCAATTCAGAGAGGTTGAACCGCTTGAAGAGGTGAAGCAATCACA	427	
Db	385	GATGACCGCCCAATTCAGAGAGGTTGAACCGCTTGAAGAGGTGAAGCAATCACA	444	
Qy	428	GTGTTAAAGAGACACGTTGAATGATGACGGCTGCTCCTATGTTGAAATTTGTTTAT	487	
Db	445	GTGTTAAAGAGACACGTTGAATGATGACGGCTGCTCCTATGTTGAAATTTGTTTAT	504	
Qy	488	TAAATTCCTCCAATAAAGCTTTACAGCCTTCTGCAAGAAAAA	538	
Db	505	TAAATTCCTCCAATAAAGCTTTACAGCCTTCTGCAAGAAAAA	555	

RESULT 2
LOCUS BG773070 813 bp mRNA linear EST 15-MAY-2001
DEFINITION 602721370F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:4838176 5',
mRNA sequence.
ACCESSION BG773070
VERSION BG773070.1 GI:14083723
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM10771 row: 1 column: 17
High quality sequence stop: 778.

FEATURES
Location/Qualifiers
1..813
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4838176"
/clone_lib="NIH_MGC_97"
/lab_host="DH10B"
/note="Organ: testis; Vector: pBluescriptR (modified pBluescript KS+); Site:1: BamHI; Site:2: SalI-xhoI (gtcgag); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTTVN-3', size-selected for average insert size 2.2 kb and normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIMH/NHGRI, National Institutes of Health). Note: this is a NIH_MGC Library."

BASE COUNT	215 a	163 c	209 g	226 t
ORIGIN				
Query Match	91.6%	Score 492.6	DB 10	Length 813
Best Local Similarity	96.8%	Pred. No. 2.8e-104		
Matches 514	Conservative 0	Mismatches 14	Indels 3	Gaps 1

Qy	1	ACGCCAGGAGCTGTGAGGAGTGCTGTGGTTCCTCGCTCCGACTCTTTTCTCTCT	60	
Db	10	ACGCCAGGAGCTGTGAGGAGTGCTGTGGTTCCTCGCTCCGACTCTTTTCTCTCT	69	
Qy	61	ACTGAGATTCATCTGTGTAATATGAGTTGCCAGGAAGATCGACCTCTTATCGGCCTA	117	
Db	70	ACTGAGATTCATCTGTGTAATATGAGTTGCCAGGAAGATCGACCTCTTATCGGCCTA	129	
Qy	118	GACCAAGAGCTACGTAGAGCTCTCTGAAATGATTTGGGCTATGCGGCCGAGCAGTTCA	177	
Db	130	GACCAAGAGCTATGTACAGCTCTCTGAAATGATTTGGGCTATGCGGCCGAGCAGTTCA	189	
Qy	178	GTGATGAAGTGGACCAACGACCACTCTGAAGAGGGGAACCACTCAACGTCAGGATC	237	
Db	190	GTGATGAAGTGGACCAACGACCACTCTGAAGAGGGGAACCACTCAACGTCAGGATC	249	
Qy	238	CTGCAGCTCTCAGGAGGGAGAGGATGAGGAGCATCTGCAGGTCAAGGCCGAGACCTG	297	
Db	250	CTGCAGCTCTCAGGAGGGAGAGGATGAGGAGCATCTGCAGGTCAAGGCCGAGACCTG	309	
Qy	298	AGCTCATAGCCAGGAACAGGGTCAACCCACAGACTGGGTGTGAGTGTGAAGTGGTCTG	357	
Db	310	AGCTCATAGCCAGGAACAGGGTCAACCCACAGACTGGGTGTGAGTGTGAAGTGGTCTG	369	
Qy	358	ATGGGCAGGAGATGGACCCCAAAATCCAGAGGAGGTCAAAACGCTGAAGAGGTGAAA	417	
Db	370	ATGGGCAGGAGATGGACCCCAAAATCCAGAGGAGGTCAAAACGCTGAAGAGGTGAAA	429	
Qy	418	AGCAATCAGAGTTTAAAGAAGACACGTTGAAATGATGCAGGCTGCTCTATGTTGGAA	477	
Db	430	GGCAATCAGAGTTTAAAGAAGACACGTTGAAATGATGCAGGCTGCTCTATGTTGGAA	489	
Qy	478	ATTGTTTCAATTAATTTCTCCCAATAAAGCTTTACAGCTTCTGCAAGAA 528		
Db	490	AATTCTTCAATTAATTTCTCCCAATAAAGCTTTACAGCTTCTGCAAGAA 540		

RESULT 3
BI826605
LOCUS 603077056F1 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:516892 5',
mRNA sequence.
ACCESSION BI826605
VERSION BI826605.1 GI:15938155

KEYWORDS EST.
SOURCE
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS 1 (bases 1 to 527)
TITLE NIH-MGC http://mgi.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11419 row: p column: 13
High quality sequence stop: 519.
Location/Qualifiers
1..527
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:516892"
/clone_lib="NIH_MGC_119"
/tissue_type="medulla"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMV-SPORT6; Site.1: NotI;
Site.2: EcoRV (destroyed); RNA source normal medulla from
anonymous male age 27. Library is oligo-dT primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.3 kb, insert size range
0.9-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 013. Note:
this is a NIH_MGC Library."
BASE COUNT 154 a 111 c 151 g 111 t
ORIGIN
Query Match 91.4%; Score 492; DB 10; Length 527;
Best Local Similarity 98.3%; Pred. No. 3.8e-104;
Matches 519; Conservative 0; Mismatches 5; Indels 4; Gaps 2;
QY 12 CTGTGAGCAGTGTGTGGTTCCTCGCGTCCGGACTCTTTTCTCTACTGAGATTCA 71
DB 1 CTGTGAGCAGTGTGTGGTTCCTCGCGTCCGGACTCTTTTCTCTACTGAGATTCA 60
QY 72 TCTGTGAAATATGAGTTGGCGAGGAAGATCGACC---TATCGGCCTAGACCAAGCGC 128
DB 61 TCTGTGAAATATGAGTTGGCGAGGAAGATCGACCCTATTATTGGCCTAGACCAAGCGC 120
QY 129 TACGTAGAGCCTCTGAAATGATGGGCGCTATGCGGCCCGAGCAGTTCAGTGAAGTGTG 188
DB 121 TATGTACAGCCTCTGAAATGATGGGCGCTATGCGGCCCGAGCAGTTCAGTGAAGTGTG 180
QY 189 GAACCAAGCAACCTGAGAGAGGGGAACCAAGCAACTCAAGTCAGGATTCCTGAGCTGCT 248
DB 181 GAACCAAGCAACCTGAGAGAGGGGAACCAAGCAACTCAAGTCAGGATTCCTGAGCTGCT 240
QY 249 CAGGAGGAGAGGATGAGGAGCATCTGAGGTCAGGCGCCGAGCCGTGAGCTCATAGC 308
DB 241 CAGGAGGAGAGGATGAGGAGCATCTGAGGTCAGGCGCCGAGCCGTGAGCTCATAGC 300
QY 309 CAGGAACAGGGTCAACCCACAGACTGGGTGTGAGTGTGAAGATGGTCTGTATGGGCGAGG 368
DB 301 CAGGAACAGGGTCAACCCACAGACTGGGTGTGAGTGTGAAGATGGTCTGTATGGGCGAGG 360
QY 369 ATGACCCGCCCAATCCAGAGAGAGGTGAAACGCCCTGAGAGAGGTGAAACATACAG 428
DB 361 ATGACCCGCCCAATCCAGAGAGAGGTGAAACGCCCTGAGAGAGGTGAAACATACAG 420

QY 429 TGTAAAAGACACACGCTTGAAATGATGACGCTGCTCTCTATTTGGAAATTTGTTCAAT 488
DB 421 TGTAAAAGACGACGCTTGAAATGATGACGCTGCTCTCTATTTGGAAATTTGTTCAAT 480
QY 489 AAAATTCCTCCCAATAAGCTTTACAGCCTTCTGCAAGAAAAA 536
DB 481 AAAATTCCTCCCAATAAGCTTTACAG-CTTCTGCAAGAAAAA 527
RESULT 4
BG250953 602363801F1 NIH_MGC_90 851 bp mRNA linear EST 13-FEB-2001
DEFINITION mRNA sequence.
ACCESSION BG250953
VERSION BG250953.1 GI:12760769
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 851)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10292 row: b column: 02
High quality sequence stop: 531.
Location/Qualifiers
1..851
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4471969"
/clone_lib="NIH_MGC_90"
/tissue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: liver; Vector: pCMV-SPORT6; Site.1: NotI;
Site.2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.7 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
BASE COUNT 224 a 205 c 239 g 183 t
ORIGIN
Query Match 91.4%; Score 492; DB 10; Length 851;
Best Local Similarity 96.4%; Pred. No. 3.9e-104;
Matches 515; Conservative 0; Mismatches 15; Indels 4; Gaps 1;
QY 8 GGAGCTGTGAGCAGTGTGTGGTTCCTCGCGTCCGGACTCTTTTCTCTACTGAGA 67
DB 11 GGAGCTGTGAGCAGTGTGTGGTTCCTCGCGTCCGGACTCTTTTCTCTACTGAGA 70
QY 68 TTCATCTGTGAAATATGAGTTGGCGAGGAAGATCGACCTA---TCGGCCTAGACCAA 123
DB 71 TTCATCTGTGAAATATGAGTTGGCGAGGAAGATCGACCTATTATTGGCCTAGACCAA 130
QY 124 GACCTAGCTAGAGCCTCTGAAATGATTTGGGCTATGCGGCCCGAGCAGTTCACTGATG 183
DB 131 GGCCTATGTCAGCCTCTCTGAAGTGGGCTATGCGGCCCGAGCAGTTCACTGATG 190
QY 184 AAGTGGAAACCAACACCTGAAAGAGGGGAACCAAGCAACTCAACGTCAGGATCTGCAG 243
DB 191 AAGTGGAAACCAACACCTGAAAGAGGGGAACCAAGCAACTCAACGTCAGGATCTGCAG 250

```
Qy 244 CTGCTCAGGAGGAGGAGGAGCATCTGCAGGTCAGAGGCCGAGCCCTGAAGCTC 303
|||||
Db 251 CTGCTCAGGAGGAGGAGGAGCATCTGCAGGTCAGAGGCCGAGCCCTGAAGCTG 310
|||||
Qy 304 ATAGCCAGGAACAGGGTCACCCACAGACTGGGTGTGAGTGTGAAGATGGTCTGTATGGGC 363
|||||
Db 311 ATAGCCAGGAACAGGGTCACCCACAGACTGGGTGTGAGTGTGAAGATGGTCTGTATGGGC 370
|||||
Qy 364 AGGAGATGACCGCCCAATCCAGAGGAGTGAACCCCTCAAGAGGTGAAGCAAT 423
|||||
Db 371 AGGAGATGACCGCCCAATCCAGAGGAGTGAACCCCTCAAGAGGTGAAGCAAT 430
|||||
Qy 424 CACAGTGTAAAGAGACACGTTGAAATGATGAGGCTGTCTCTATGTTGAAATTTGT 483
|||||
Db 431 CACAGTGTAAAGAGACGTTGAAATGATGAGGCTGTCTCTATGTTGAAATTTGT 490
|||||
Qy 484 TCATTAAATTTCTCCCAATAAAGCTTTACAGCCTTCTCCAAAGAAAAA 537
|||||
Db 491 TCATTAAATTTCTCCCAATAAAGCTTTACAGCCTTCTCCAAAGAAAAA 544
|||||

RESULT 5
BG120838 572 bp mRNA linear EST 30-JAN-2001
LOCUS 602351010F1 NIH_MGC_90 Homo sapiens cDNA clone IMAGE:4445882 5',
DEFINITION mRNA sequence.
ACCESSION BG120838
VERSION BG120838.1 GI:12614347
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10224 row: c column: 03
High quality sequence start: 5
High quality sequence stop: 572.
Location/Qualifiers
1. .572
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4445882"
/clone_lib="NIH_MGC_90"
/tissue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: liver; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.7 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

BASE COUNT 167 a 124 c 166 g 115 t
ORIGIN

Query Match 91.2%; Score 490.6; DB 10; Length 572;
Best Local Similarity 96.7%; Pred. No. 8.1e-104;
Matches 523; Conservative 0; Mismatches 14; Indels 4; Gaps 2;

Qy 1 ACCCCAGGAGCTGTGAGGCAGTGTGTGGTTCTCTGCGTCGCGACTCTTTTCTCT 60
|||||
Db 28 ACCCAGGAGCTGTGAGGCAGTGTGTGGTTCTCTGCGTCGCGACTCTTTTCTCT 87
|||||
```

```
Qy 61 ACTGAGATTCATCTGTGTGAAATATGAGTTGGCGAGGAAGATCGACC---TATCGCCTA 117
|||||
Db 88 ACTGAGATTCATCTGTGTGAAATATGAGTTGGCGAGGAAGATCGACCTATTATTGGCCTA 147
|||||
Qy 118 GACCAAGACGCTACGTAGAGCCCTCCTGAAATGATTTGGGCTATGCGGCCCGAGCAGTTCA 177
|||||
Db 148 GACCAAGGCGCTATGTACAGCCTCCTGAAGTSAATTGGGCTATGCGGCCCGAGCAGTTCA 207
|||||
Qy 178 CTGATCAACTGGAACCCAGCAACACCTGAAGAAGGSGAACCCAGCAACTCAACGTCAGGATC 237
|||||
Db 208 GTGATGAAGTGGAACCCAGCAACACCTGAAGAAGGSGAACCCAGCAACTCAACGTCAGGATC 267
|||||
Qy 238 CTGCGAGCTGCTCAGGAGGAGAGGATGAGGAGGATCTGCAGGTCAAGGCCCGAAGCCTG 297
|||||
Db 268 CTGCGAGCTGCTCAGGAGGAGAGGATGAGGAGGATCTGCAGGTCAAGGCCCGAAGCCTG 327
|||||
Qy 298 AAGCTCATAGCCAGGAACACAGGGTCAACCACAGACTGGGTGTGAGTGTGAAGATGGTCTG 357
|||||
Db 328 AAGCTCATAGCCAGGAACACAGGGTCAACCACAGACTGGGTGTGAGTGTGAAGATGGTCTG 387
|||||
Qy 358 ATGGGCAGGAGATGGACCCGCCCAATCCAGAGGAGGTGAAAAACGCTGAAGAGGTGAAA 417
|||||
Db 388 ATGGGCAGGAGATGGACCCGCCCAATCCAGAGGAGGTGAAAAACGCTGAAGAGGTGAAA 447
|||||
Qy 418 AGCAATCACAGTGTAAAAAGAACACAGCTTGAATGATGACGAGCTCTCTATGTTGGA 477
|||||
Db 448 AGCAATCACAGTGTAAAAAGAACACAGCTTGAATGATGACGAGCTCTCTATGTTGGA 506
|||||
Qy 478 ATTTGTTCAATAAAATTTCTCCCAATAAAGCTTTACAGCTTCTGCAAGAAAAA 537
|||||
Db 507 ATTTGTTCAATAAAATTTCTCCCAATAAAGCTTTACAGCTTCTGCAAGAAAAA 566
|||||
Qy 538 A 538
Db 567 A 567

RESULT 6
BG18671 509 bp mRNA linear EST 11-OCT-2001
LOCUS 603392594F1 NIH_MGC_90 Homo sapiens cDNA clone IMAGE:5402663 5',
DEFINITION mRNA sequence.
ACCESSION BG18671
VERSION BG18671.1 GI:16042344
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12026 row: 1 column: 24
High quality sequence stop: 509.
Location/Qualifiers
1. .509
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5402663"
/clone_lib="NIH_MGC_90"
/tissue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"

FEATURES
source
```


/note="Organ: liver; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally; oligo-dr primed. Average insert size 1.7 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."

BASE COUNT	142 a	107 c	154 g	106 t
ORIGIN				
Query Match	88.6%;	Score 476.8;	DB 10;	Length 509;
Best Local Similarity	98.0%;	Pred. No. 1.3e-100;		
Matches 494;	Conservative 0;	Mismatches 7;	Indels 3;	Gaps 1;
Qy	1	ACGCCAGGAGCTGTGAGGAGCTGCTGTGTGGTTCCTGCCGCTCCGGACTCTTTTTCCTCT	60	
Db	6	ACGCTAGGAGAGCTGTGAGGAGCTGCTGTGTGGTTCCTGCCGCTCCGGACTCTTTTTCCTCT	65	
Qy	61	ACTGAGATTCATCTGTGTGAATATGAGTTGGCAGAGNAGATCGACC---	117	
Db	66	ACTGAGATTCATCTGTGTGAATATGAGTTGGCAGAGNAGATCGACCTATTATTGGCCTA	125	
Qy	118	GACCAAGACGCTACGTAGAGCCTCTGAAATGATTGGCCCTATGCGGCCCGAGCACTTCA	177	
Db	126	GACCAAGCGCTATGTACAGCCTCTGAATGATTGGCCCTATGCGGCCCGAGCACTTCA	185	
Qy	178	GTGATGAAGTGGAAACCAACACCTGAAGAAGGGGAACCAACCACTCAACGCTCAGGATC	237	
Db	186	GTGATGAAGTGGAAACCAACACCTGAAGAAGGGGAACCAACCACTCAACGCTCAGGATC	245	
Qy	238	CTGCAGCTGCTCAGGAGGAGAGATGAGGAGCATCTGCAGGTCAAGGCCGGAAGCCCTG	297	
Db	246	CTGCAGCTGCTCAGGAGGAGAGATGAGGAGCATCTGCAGGTCAAGGCCGGAAGCCCTG	305	
Qy	298	AAGCTCATAGCCAGGAACAGGCTCACCCACAGACTGGGTGTGAGTGTGAAGATGTGCTCTG	357	
Db	306	AAGCTCATAGCCAGGAACAGGCTCACCCACAGACTGGGTGTGAGTGTGAAGATGTGCTCTG	365	
Qy	358	ATGGCAGGAGATGGACCCGCAATCCAGAGAGGTGAAACCGCTCGAAGAAGGTGAAA	417	
Db	366	ATGGCAGGAGATGGACCCGCAATCCAGAGAGGTGAAACCGCTCGAAGAAGGTGAAA	425	
Qy	418	AGCAATCACAGTGTAAAAGAACACGCTTGAATGATGCAGGCTGCTCTATGTTGGAA	477	
Db	426	AGCAATCACAGTGTAAAAGAACACGCTTGAATGATGCAGGCTGCTCTATGTTGGAA	485	
Qy	478	ATTGTTTCATTAATAATTTCTCCCAA	501	
Db	486	ATTGTTTCATTAATAATTTCTCCCAA	509	

RESULT 7	
BMI72302	
LOCUS	623 bp mRNA linear EST 04-DEC-2001
DEFINITION	ImageQC_4_2001/smul188bdf41.x1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:4838176 5', mRNA sequence.
ACCESSION	BMI72302
VERSION	BMI72302.1
KEYWORDS	EST.
SOURCE	GI:17311865
ORGANISM	human.
REFERENCE	Homo sapiens
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE	1 (bases 1 to 623)
JOURNAL	Kale, P.I., Harsch, T.J., Folta, P.A., Nelson, D.O., Sanders, C.G. and Prange, C.K.
COMMENT	The I.M.A.G.E. Consortium quality control effort: clone unpublished (2001)
	Other_ESTs: BG773070
	Contact: Prange CK
	The I.M.A.G.E. Consortium
	Lawrence Livermore National Laboratory
	Livermore, CA, USA

Email: help@image.llnl.gov
This read has been verified (found to hit its original self in the correct orientation), as part of the I.M.A.G.E. Consortium quality control effort. High quality sequence is defined as having 100 or more base pairs with a phred quality value of 20 or greater, where a sliding window of 4 base pairs with a phred quality value of 15 or greater marks the beginning and end of the sequence. For information on obtaining this clone, please contact
help@image.llnl.gov.

Plate: LLAM10771 row: 1 column: 17
Seq primer: -21ml3
High quality sequence stop: 623.
Location/Qualifiers

FEATURES
source

1. 623
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4838176"
/clone_lib="NIH_MGC_97"
/lab_host="DH10B"
/note="Organ: testis; Vector: pBluescriptR (modified pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgaag); Oligo-dr primed using primer 5'-TTTTTTTTTTTTTTTTVN-3', size-selected for average insert size 2.2 kb and normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Garninci, in preparation). Library constructed by M. Brownstein (NIMH/NHGRI, National Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT 166 a 125 c 173 g 159 t
ORIGIN

Query Match 86.1%; Score 463.4; DB 10; Length 623;
Best Local Similarity 94.5%; Pred. No. 1.7e-97;
Matches 502; Conservative 0; Mismatches 26; Indels 3; Gaps 2;

Qy	1	ACGCCAGGAGCTGTGAGGAGCTGCTGTGTGGTTCCTGCCGCTCCGGACTCTTTTTCCTCT	60	
Db	10	ACGCCAGGAGCTGTGAGGAGCTGCTGTGTGGTTCCTGCCGCTCCGGACTCTTTTTCCTCT	69	
Qy	61	ACTGAGATTCATCTGTGTGAATATGAGTTGGCAGAGNAGATCGACCTAT--CGGCCCTAG	118	
Db	70	ACTGAGATTCATCTGTGTGAATATGAGTTGGCAGAGNAGATCGACCTATATTATGGCCTA	129	
Qy	119	ACCAAGAGCTACGTAGAGCCTCTGAAATGATTGGCCCTATGCGGCCCGAGCAGTTCAG	178	
Db	130	TATCATGCTCTATGTACAGCCTCTGAAATGATTGGGCCCTATGCGGCCCGAGCAGTTCAG	189	
Qy	179	TGATGAAGTGGAAACCAACACCTGAAGAAGGGGAACCAACCACTCAACGTCAGGATCC	238	
Db	190	TGATGAAGTGGAAACCAACACCTGAAGAAGGGGAACCAACCACTCAACGTCAGGATCC	249	
Qy	239	TGCAGCTGCTCAGGAGGAGAGATGAGGAGCATCTGCAGGTCAAGGCCCGAAGCCCTGA	298	
Db	250	TGCAGCTGCTCAGGAGGAGAGATGAGGAGCATCTGCAGGTCAAGGCCCGAAGCCCTGA	309	
Qy	299	AGCTCATAGCCAGGAGGAGGTCACCCACAGACTGGGTGTGAGTGTGAAGATGGTCTCTGA	358	
Db	310	AGCTCATAGCCAGGAGGAGGTCACCCACAGACTGGGTGTGAGTGTGAAGATGGTCTCTGA	369	
Qy	359	TGGGCAGGAGATGGACCCGCAATCCAGAGAGGTGAAACCGCTCGAAGAAGGTGAAAA	418	
Db	370	TGGGCAGGAGATGGACCCGCAATCCAGAGAGGTGAAACCGCTCGAAGAAGGTGAAAGG	429	
Qy	419	GCAATCACAGTGTAAAAGAACACGTTGAAATGATGCAGGCTGCTCTATGTTGGAAA	478	
Db	430	GCAATCACAGTGTAAAAGAACACATGCTGAAATGTTGCAGGCTGCTCTATGTTGGAAA	489	
Qy	479	TTTGTTCATTAATAATTTCTCCCAA	528	
Db	490	ATTCTTCATTGAAGTTCTCCCAATAAAGCTTTTACAGCCTTCTGANAAGAA	540	

QY 134 AGAGCCTCCTGAAATGATTGGCCCTATCGGCCCGAGCGATTCAGTCAAGTGGAAACC 193
Db 397 ACAGCCTCCTGAAATGATTGGCCCTATCGGCCCGAGCGATTCAGTCAAGTGGAAACC 338
QY 194 AGCAACACCTGAAGAAGGGGAACAGCAACCTCAACGTCAGGATCCCTGCAGTCTCAGGA 253
Db 337 AGCAACACCTGAAGAAGGGGAACAGCAACCTCAACGTCAGGATCCCTGCAGTCTCAGGA 278
QY 254 GGGAGAGGATGAGGAGGATCTGCAGGTCAGAGGCGCAAGGCGCAAGCCTGAAGCTCATAGCCAGGA 313
Db 277 GGGAGAGGATGAGGAGGATCTGCAGGTCAGAGGCGCAAGGCGCAAGCCTGAAGCTCATAGCCAGGA 218
QY 314 ACAGGTCACCCACAGACTGGGTGAGTGCAAGATGCTCTGATGGGAGAGATGGA 373
Db 217 ACAGGTCACCCACAGACTGGGTGAGTGCAAGATGCTCTGATGGGAGAGATGGA 158
QY 374 CCGCCCAATCCACAGGAGGTGAAAACGCTCAAGAAAGGTGAAAACCAATCACAGTGTTA 433
Db 157 CCGCCCAATCCACAGGAGGTGAAAACGCTCAAGAAAGGTGAAAACCAATCACAGTGTTA 98
QY 434 AAGAAGACACGTTGAAATGATGAGGCTGCTCTATGTTGGAAATTTGTTCAATAAAT 493
Db 97 AAGAAGGACGTTGAAATGATGAGGCTGCTCTATGTTGGAAATTTGTTCAATAAAT 38
QY 494 TCTCCCAATAAAGCTTTACAGCCTTCTGCAAGAAAA 530
Db 37 TCTCCCAATAAAGCTTTACAGCCTTCTGCAAGAAAA 1

RESULT 10
AI381509/c
LOCUS
DEFINITION
te76b07.x1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone
IMAGE:2092597 3' similar to TR:Q13070 Q13070 GAGE-6 PROTEIN. ;
mRNA sequence.
AI381509
AI381509.1 GI:4194290
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 447)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 497 Std Error: 0.00
Seq primer: -40UP from Gibco.
Location/Qualifiers
1. .447
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2092597"
/clone.lib="Soares_NFL_T_GBC_S1"
/lab_host="DH10B"
/note="Organ: pooled; Vector: p7T3D-Pac (Pharmacia) with
a modified polylinker; Site_1: Not 1; Site_2: Eco RI;
Equal amounts of plasmid DNA from three normalized
libraries (fetal lung NbHL19w, testis NHT, and B-cell
NCI-CGAP-GCB1) were mixed, and ss circles were made in
vitro. Following HAP purification, this DNA was used as
tracer in a subtractive hybridization reaction. The driver
was PCR-amplified cDNAs from pools of 5,000 clones made
from the same 3 libraries. The pools consisted of
I.M.A.G.E. clones 297480-302087, 682632-687239,
726408-728711, and 729096-731399. Subtraction by Bento
Soares and M. Fatima Bonaldo."

BASE COUNT 84 a 130 c 95 g 138 t
ORIGIN
Query Match 78.9%; Score 424.6; DB 9; Length 447;
Best Local Similarity 98.4%; Pred. No. 1.7e-88;
Matches 440; Conservative 0; Mismatches 4; Indels 3; Gaps 1;
QY 88 GTTGGCGAGGAAGATCGACC---TATCGGCCTAGACCAAGCCTAGCTAGAGCCTCCTG 144
Db 447 GTTGGCGAGGAAGATCGACCCTATTATTGGCTTAGACCAAGGCGCTATGTACAGCCTCCTG 388
QY 145 AAATGATTTGGCCCTATGCGGCCCGAGCAGCTTCAGTCAAGTGGAAACCAACACCTG 204
Db 387 AAATGATTTGGCCCTATGCGGCCCGAGCAGCTTCAGTCAAGTGGAAACCAACACCTG 328
QY 205 AAGAAGGGGAACCAAGCAACTCAACGTCAGGATCCCTGAGTGGGAGGAGAGGATG 264
Db 327 AAGAAGGGGAACCAAGCAACTCAACGTCAGGATCCCTGAGTGGGAGGAGAGGATG 268
QY 265 AGGAGGATCTGCAAGTCAAGGGCGCAAGCCTGAAGCTCATAGCCAGGAACAGGGTCAAC 324
Db 267 AGGAGGATCTGCAAGTCAAGGGCGCAAGCCTGAAGCTCATAGCCAGGAACAGGGTCAAC 208
QY 325 CACAGACTGGTGTGAGTGTGAAGATGGTCTCATGGCAGGAGATGGACCCGCCAATC 384
Db 207 CACAGACTGGTGTGAGTGTGAAGATGGTCTCATGGCAGGAGATGGACCCGCCAATC 148
QY 385 CAGAGGAGGTGAAAACGCTGAAGAAAGGTGAAAAGCAATCACAGTGTGTTAAAGAGACAC 444
Db 147 CAGAGGAGGTGAAAACGCTGAAGAAAGGTGAAAAGCAATCACAGTGTGTTAAAGAGACAC 88
QY 445 GTTGAATGATGCGAGGCTGCTCTATGTTGGAAATTTGTTCAATAAATTTCCCAATAA 504
Db 87 GTTGAATGATGCGAGGCTGCTCTATGTTGGAAATTTGTTCAATAAATTTCCCAATAA 28
QY 505 AGCTTTACAGCCTTCTGCAAGAAAAA 531
Db 27 AGCTTTACAGCCTTCTGCAAGAAAAA 1

RESULT 11
AA738037/c
LOCUS
DEFINITION
nx15el1.sl NCI-CGAP_GC3 Homo sapiens cDNA clone IMAGE:1256204 3'
similar to TR:Q13070 Q13070 GAGE-6 PROTEIN. ; mRNA sequence.
AA738037
AA738037.1 GI:2768794
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 464)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/dbbrp/image/image.html
Insert Length: 511 Std Error: 0.00
Seq primer: -40ml3 fwd. Et from Amersham
High quality sequence stop: 435.
Location/Qualifiers
1. .464
source

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1256204"
/clone_lib="NCI_CGAP_GC3"
/tissue_type="pooled germ cell tumors"
/lab_host="DH10B"
/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; 1st strand cDNA was prepared from 3 pooled germ cell tumors, and was then primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library is not normalized. Library was constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT 86 a 132 c 93 g 153 t
ORIGIN

Query Match 78.9%; Score 424.4; DB 9; Length 464;
Best Local Similarity 96.9%; Pred. No. 1.9e-88;
Matches 444; Conservative 0; Mismatches 11; Indels 3; Gaps 1;
QY 84 ATGAGTTGGCAGGAGATCGACC---TATCGGCTAGACCAAGCGTACGTAGAGCCT 140
|||||
Db 464 ATGAGTTGGCAGGAGATCGACCTGTATTGGCCTAGTCCAGACGCTATGTACAGCCT 405
|||||
QY 141 CCTGAAATGATTGGGCTATGCGCCCGAGCAGTTCAAGTGAAGTGAACCAACCAACA 200
|||||
Db 404 CCTGAAATGATTGGGCTATGACGCGGAGCAGTTCAAGTGAAGTGAACCAACCAACA 345
|||||
QY 201 CCTGAAGAGGGGAACCACTCAACGTCAGGATCCTGCAGCTGCTCAGGAGGGAGAG 260
|||||
Db 344 CCTGAAGAGGGGAACCACTCAACGTCAGGATCCTGCAGCTGCTCAGGAGGGAGAG 285
|||||
QY 261 GATGAGGAGCATCTCAGGTCAGGCGGAGCGGAGCCTCAAGCTCATAGCCAGGAACAGG 320
|||||
Db 284 GATGAGGAGCATCTCAGGTCAGGCGGAGCGGAGCCTCAAGCTCATAGCCAGGAACAGG 225
|||||
QY 321 CACCACAGACTGGGTGTGAGTGTGAAGATGGTCTGATGGCGAGGAGATGGACCCGCCA 380
Db 224 CACCACAGACTGGGTGTGAGTGTGAAGATGGTCTGATGGCGAGGAGATGGACCCGCCA 165
QY 381 AATCCAGAGGAGTTGAAACGCTGAAGAGGTTGAAAGCAATCACAGTGTAAAGAG 440
Db 164 AATCCAGAGGAGTTGAAACGCTGAAGAGGTTGAAAGCAATCACAGTGTAAAGAG 105
QY 441 ACAGTTGAAATGATGCAGGCTGCTCTATGTTGGAAATTGTTCAATTAATAATCTCCCA 500
Db 104 GCAGTTCAATGATGCAGGCTGCTCTATGTTGGAAATTGTTCAATTAATAATCTCCCA 45
QY 501 ATAAAGCTTTACAGCCTTCTGCAAGAAAGAAAAA 538
Db 44 ATAAAGCTTTACAGCCTTCTGCAAGAAAGAAAAA 7

RESULT 12
BG700088
LOCUS 581 bp mRNA linear EST 07-MAY-2001
DEFINITION 602679431F1 NIH_MGC_95 Homo sapiens cDNA clone IMAGE:4812462 5', mRNA sequence.
ACCESSION BG700088
VERSION BG700088.1 GI:13969078
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 581)
REFERENCE NIH-MGC <http://mgc.nci.nih.gov/>.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLNL0704 row: m column: 07
High quality sequence stop: 548.
Location/Qualifiers
1. 581
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4812462"
/clone_lib="NIH_MGC_95"
/tissue_type="hippocampus"
/lab_host="DH10B"
/note="Organ: brain; Vector: pBluescriptR (modified pBluescript KS+); Site.1: BamHI; Site.2: SalI-XhoI (gtcgaag); Oligo-dT primed using primer 5'-gtttttttttttttttttn-3', size-selected for average insert size 2.5 kb and normalized to 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NHGRI, National Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT 164 a 142 c 166 g 109 t
ORIGIN

Query Match 78.8%; Score 424; DB 10; Length 581;
Best Local Similarity 92.8%; Pred. No. 2.4e-88;
Matches 501; Conservative 0; Mismatches 30; Indels 9; Gaps 5;
QY 8 GGAGCTGTGAGCAGTGCTGTGGTTCCTGCGCTCGGACTCTCTTTCTCTACTGAGA 67
Db 25 GGAGCTGTGAGCAGTGCTGTGGTTCCTGCGCTCGGACTCTCTTTCTCTACTGAGA 84
QY 68 TTCATCTGTGAAATATGATGTTGGCGAGGAAGATGCACCTATCGCCCTAGACCAGAGC 127
Db 85 TTCATCTGTGAAATATGATGTTGGCGAGGAAGATGCACCTATCGCCCTAGACCAGAGC 144
QY 128 CTAGCTAGAGCCTCTCTGA-AATGATTGGGCTTATGCGGCCCGCAGCAGTTCAAGTGAAG 186
Db 145 CTAGCTAGAGCCTCTCTGACAATGATTGGGCTTATGCGGCCCGCAGCAGTTCAAGTGAAG 204
QY 187 TGGAACCAACACACTGAGAGAGGGGAACCACTCAACGTCAGGATCCTGCAGCTG 246
Db 205 TGGAACCAACACACTGAGAGAGGGGAACCACTCAACGTCAGGATCCTGCAGCTG 264
QY 247 CTCAGGAGGGAG-AGGATGAGGAGCATCTGCAGGTCAAGGCGCGAAGCCTGAAGTCTAT 305
Db 265 CTCAGGAGGGAGCGCTGAGGAGCATCTGCAGGTCAAGGCGCGAAGCCTGAAGTCTAT 324
QY 306 AGCCAGGAACAGGGTCAACCCACAGACTGGGTGTGAGTGTGAAGATGGTCTGTATGGGCGAG 365
Db 325 AGCCAGGAACAGGGTCAACCCACAGACTGGGTGTGAGTGTGAAGATGGTCTGTATGGGCGAG 384
QY 366 GAGATGGACCGCCCAATCCAGAGGAGTGAACACGCTGAAGA-AGCTGAAAGCAATC 424
Db 385 GAGATGGACCGCCCAATCCAGAGGAGTGAACACGCTGAAGACAGGTGAACAGCAATC 444
QY 425 ACA---GTGTTAAAGAAGACACAGTTGAAATGATGACAGGCTGCTCTCTATGTTGGAATTT 481
Db 445 ACAAGTGTACAAAGAAAGGACCGTTGACATGATGACAGGCTGCTCTATGTTGGAATCTC 504
QY 482 GTTCA---TTAAATTCCTCCCAATAAAGCTTTACAGCCTTCTGCAAGCAAGAAAAA 538
Db 505 TGTAACATTACACATTTCTCCCACTAAAGCTATACAGCCTTCTGCAACAAAAACACACAAA 564
RESULT 13

AW016546/c
LOCUS AW016546 418 bp mRNA linear EST 10-SEP-1999
DEFINITION UI-H-BiOp-abg-g-06-0-UI.s1 NCI_CGAP_Sub2 Homo sapiens cDNA clone
IMAGE:2711986 3', mRNA sequence.
ACCESSION AW016546
VERSION AW016546.1 GI:5865303
KEYWORDS EST
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 418)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
The sequence contained an oligo-dT track that was present in the
oligonucleotide that was used to prime the synthesis of first
strand cDNA and therefore this may represent a bonafide poly A
tail. cDNA Library Preparation: M.B. Soares Lab Clone Distribution:
NCI-CGAP clone distribution information can be found through the
I.M.A.G.E. Consortium/LLNL at:
[www-bio.llnl.gov/bbrp/image/image.html](http://www.bio.llnl.gov/bbrp/image/image.html)
Seq primer: M13 Forward
POLYA-Yes.

FEATURES

source

Location/Qualifiers

1. 418
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2711986"
/clone_lib="NCI_CGAP_Sub2"
/lab_host="DH10B (Life Technologies)"
/note="vector: p7T3D-Pac (Pharmacia) with a modified
polylinker. Site.1: Not I; Site.2: Eco RI; The
NCI_CGAP_Sub2 library is a substracted library derived from
BI. BI constitutes a mixture of 21 normalized or
substracted NCI_CGAP libraries: NCI_CGAP_Co4, NCI_CGAP_Pr22
, NCI_CGAP_Pr28, NCI_CGAP_Co10, NCI_CGAP_Co16,
NCI_CGAP_Kid5, NCI_CGAP_Kid12, NCI_CGAP_Kid3,
NCI_CGAP_Kid11, NCI_CGAP_Lym2, NCI_CGAP_Br2, NCI_CGAP_Lu5
NCI_CGAP_CUL1, NCI_CGAP_Le12, NCI_CGAP_Brn23, NCI_CGAP_GC6
, NCI_CGAP_Lu24, NCI_CGAP_Lu19, NCI_CGAP_GC4, NCI_CGAP_GC6
, NCI_CGAP_Brn25. These 21 libraries were pooled and a
single-stranded DNA preparation of the resulting mixture
was used as a tracer in a subtractive hybridization with a
driver whose composition is detailed below: NCI_CGAP_Kid3
pool 1 LLAM 3334-3337, 3682-3683, 3798-3803 (IMAGE)
CloneIDs 1322376-1323911, 1456008-1456775, 1500552-1502855
) NCI_CGAP_Kid5 pool 1 LLAM 3338-3342, 3722-3725,
3776-3778 (IMAGE CloneIDs 1323912-1325831,
1471368-1472903, 1492104-1493255) NCI_CGAP_Lu5 pool 1 LLAM
3575-3582, 3851-3854 (IMAGE CloneIDs 1414920-1417991,
1520904-1522439) NCI_CGAP_GC4 pool 1 LLAM 3164-3167,
3716-3720, 3733-3735 (IMAGE CloneIDs 1257096-1258631,
1469064-1470983, 1475592-1476743) NCI_CGAP_Pr22 pool 1
LLAM 2457-2459, 2758-2759, 3062-3068 (IMAGE CloneIDs
983608-986759, 1101192-1101959, 1217928-1220615)
NCI_CGAP_Co10 pool 1 LLAM 2644-2653, 2871-2872 (IMAGE)
CloneIDs 1057416-1061255, 1144584-1145351) The resulting
subtracted library contained 4 million recombinants.
Subtraction was performed as previously described (Bonaldi
, Lennon & Soares (1996): Normalization and Subtraction:
Two Approaches To Facilitate Gene Discovery. Genome
Research 6, 791-806.
TAG_LIB=NCI_CGAP_GC4
TAG_TISSUE=germ cell
TAG_SEQ=AAATC"

BASE COUNT
ORIGIN

74 a 119 c 87 g 138 t

Query Match 76.9%; Score 413.8; DB 9; Length 418;
Best Local Similarity 99.5%; Pred. No. 5.5e-86;
Matches 415; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 122 AAGAGCGCTACGTAGAGCCCTCCTCAAAATGATTTGGCCCTATGCGGCCGAGCAGTTCAAGTGA 181
Db 418 AAGAGCGCTACGTAGAGCCCTCCTCAAAATGATTTGGCCCTATGCGGCCGAGCAGTTCAAGTGA 359
QY 182 TGAAGTGAACACAGCAACACCTGAAGAGGGGAACCACTCAACCTGTCAGGATCCTCTGC 241
Db 358 TGAAGTGAACACAGCAACACCTGAAGAGGGGAACCACTCAACCTGTCAGGATCCTCTGC 299
QY 242 AGCTCTCAGGAGGGAGAGATGAGGAGCATCTGCAAGGTCAAGGCCGCAAGCCTGAAGC 301
Db 298 AGCTCTCAGGAGGGAGAGATGAGGAGCATCTGCAAGGTCAAGGCCGCAAGCCTGAAGC 239
QY 302 TCATAGCCAGGAACAGGGTCAACCCACAGACTGGGTGTGAGTGTGAAGATGGTCTCTGATGG 361
Db 238 TCATAGCCAGGAACAGGGTCAACCCACAGACTGGGTGTGAGTGTGAAGATGGTCTCTGATGG 179
QY 362 GCAGGAGATGGACCCGCCCAATCCAGAGGAGGTGAAACCGCTCAAGAGGTGAAAGCA 421
Db 178 GCAGGAGATGGACCCGCCCAATCCAGAGGAGGTGAAACCGCTCAAGAGGTGAAAGCA 119
QY 422 ATCACAGTGTAAAAGAACACAGCTTGAATATGATGCAAGGCTGCTCTATGTTGGAAATTT 481
Db 118 ATCACAGTGTAAAAGAACAGGCACGTTGAAATATGATGCAAGGCTGCTCTATGTTGGAAATTT 59
QY 482 GTTCATTAAATCTCCCAATAAAGCTTTACAGCCTTCTGCAAGAGAAAAA 538
Db 58 GTTCATTAAATCTCCCAATAAAGCTTTACAGCCTTCTGCAAGAGAAAAA 2
RESULT 14
BG120336
LOCUS BG120336 480 bp mRNA linear EST 30-JAN-2001
DEFINITION 602353732F1 NIH_MGC_90 Homo sapiens cDNA clone IMAGE:4451958 5',
mRNA sequence.
ACCESSION BG120336
VERSION BG120336.1 GI:126113845
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 480)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM10239 row: p column: 07
High quality sequence stop: 480.
Location/Qualifiers
1. 480
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4451958"
/clone_lib="NIH_MGC_90"
/tissue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: liver; Vector: pCMV-SPORT6; Site.1: NotI;
Site.2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.7 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

source		1. .426			
		/organism="Homo sapiens"			
		/db_xref="taxon:9606"			
		/clone="IMAGE:1409245"			
		/clone_lib="Soares_testis_NHT"			
		/sex="male"			
		/lab_host="DH10B"			
		/note="Vector: pT73D-Pac (Pharmacia) with a modified			
		polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA			
		was prepared from mRNA obtained from Clontech Laboratories			
		, inc., and primed with a Not I - oligo(dT) primer [5',			
		TGTTACCAATCTGAAGTGGAGCGGCCCAATTTTTTTTTTTT 3'].			
		Double-stranded cDNA was ligated to Eco RI adaptors			
		(Pharmacia), digested with Not I and cloned into the Not I			
		and Eco RI sites of the modified pT73 vector. Library			
		went through one round of normalization to Cot5, and was			
		constructed by Bento Soares and M. Fatima Bonaldo."			
BASE COUNT		79 a	119 c	95 g	131 t
ORIGIN					
		Query Match			
		Best Local Similarity			
		Matches 417; Conservative			
		Score 412.8; DB 9; Length 426;			
		Pred. No. 9.4e-86;			
		Mismatches 9; Indels 0; Gaps 0;			
QY		102	TCGACCTATCGGCCTAGACCAAGCGCTACGTAGAGCCTCTCTGAAATGATTGGCGCTATG	161	
Db		426	TCGACCTATCGGCCTAGACCAAGCGCTACGTAGAGCCTCTCTGAAATGATTGGCGCTATG	367	
QY		162	CGGCCCGAGCAGTTTCAGTGTATGAAGTGAACCAACACCTGAAGAAGGGGAACCCAGCA	221	
Db		366	CGGCCCGAGCAGTTTCAGTGTATGAAGTGAACCAACACCTGAAGAAGGGGAACCCAGCA	307	
QY		222	ACTCAACGTCAGGATCCTCGCAGCTGCTCAGGAGGAGAGGATGAGGGAGCATCTGCAGGT	281	
Db		306	ACTCAACGTCAGGATCCTCGCAGCTGCTCAGGAGGAGAGGATGAGGGAGCATCTGCAGGT	247	
QY		282	CAAGGCCCAAGCCTGAAGCTATAGCCAGGAACAGGTCACCCACAGACTGGGTGTGAG	341	
Db		246	CAAGGCCCAAGCCTGAAGCTATAGCCAGGAACAGGTCACCCACAGACTGGGTGTGAG	187	
QY		342	TGTGAAGATGGTCTCTGATGGCAGGAGATGGACCCGCCAAATCCAGAGGAGGTGAAAAACG	401	
Db		186	TGTGAAGATGGTCTCTGATGGCAGGAGATGGACCCGCCAAATCCAGAGGAGGTGAAAAACG	127	
QY		402	CCTGAAGAAGGTGAAAAGCAATCACAGTGTGTAAAGAAAGACACGTTGAAATGATGCAGGC	461	
Db		126	CCTGAAGAAGGTGAAAAGCAATCACAGTGTGTAAAGAAAGACACGTTGAAATGATGCAGGC	67	
QY		462	TGCTCCTATGTTGGAATTTGTTCAATTAATTTCTCCCAATAAAGCTTTACAGCCTTCTG	521	
Db		66	TGCTCCTATGTTGGAATTTGTTCAATTAATTTCTCCCAATAAAGCTTTACAGCCTTCTT	7	
QY		522	CAAAAG 527		
Db		6	CAAAAA 1		
		Search completed: July 1, 2002, 06:39:18			
		Job time: 1653 sec			

BASE COUNT		170 a		93 c		133 g		84 t	
ORIGIN									
		Query Match							
		Best Local Similarity							
		Matches 420; Conservative							
		Score 413.4; DB 10; Length 480;							
		Pred. No. 6.8e-86;							
		Mismatches 11; Indels 0; Gaps 0;							
QY		108	TATCGGCCTAGACCAAGCAGCTACGTAGAGCCTCTCTGAAATGATTGGCGCTATGGCGCC	167					
Db		15	TATTGGCTAGACCAAGCGCTATGTACAGCCTCTCTGAAGTGTATGGCGCTATGGCGCC	74					
QY		168	GAGCAGTTTCAGTGTATGAAGTGAACCAACACACCTGAAGAAGGGGAACCCAGCAACTCA	227					
Db		75	GAGCAGTTTCAGTGTATGAAGTGAACCAACACACCTGAAGAAGGGGAACCCAGCAACTCA	134					
QY		228	CGTCAGGATCCTGCAGCTGCTCAGGAGGAGGAGGATGAGGGAGCATCTGCAGGTCAAGGG	287					
Db		135	CGTCAGGATCCTGCAGCTGCTCAGGAGGAGGAGGATGAGGGAGCATCTGCAGGTCAAGGG	194					
QY		288	CCGAAGCCTGAAGCTCATAGCCAGGAACAGGTCACCCACAGACTGGGTGTGAGTGA	347					
Db		195	CCGAAGCCTGAAGCTCATAGCCAGGAACAGGTCACCCACAGACTGGGTGTGAGTGA	254					
QY		348	GATGGTCTGATGGCGAGGAGATGACCCGCCAAATCCAGAGGAGGTGAAAACGCCCTGAA	407					
Db		255	GATGGTCTGATGGCGAGGAGATGACCCGCCAAATCCAGAGGAGGTGAAAACGCCCTGAA	314					
QY		408	GAAGGTGAAGAGCAATCACAGTGTGTAAAGAGACACCTTGAATGATGCAGGCTGCTCC	467					
Db		315	GAAGGTGAAGAGCAATCACAGTGTGTAAAGAGAGGACGTTGAAATGATGCAGGCTGCTCC	374					
QY		468	TATGTTGGAATTTGTTCAATTAATTTCTCCCAATAAAGCTTTACAGCCTTCTGCAAGA	527					
Db		375	TATGTTGGAATTTGTTCAATTAATTTCTCCCAATAAAGCTTTACAGCCTTCTGCAAAA	434					
QY		528	AAAAA	538					
Db		435	AAAAA	445					
		RESULT 15							
		AA868226/c							
LOCUS		AA868226 426 bp mRNA linear EST 16-MAR-1998							
DEFINITION		ak48h07.s1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:1409245							
		3' similar to SW:GGE2_HUMAN Q13066 GAGE-2 PROTEIN. [1] ;, mRNA							
		sequence.							
ACCESSION		AA868226							
VERSION		AA868226.1 GI:2963671							
KEYWORDS		EST.							
SOURCE		human.							
ORGANISM		Homo sapiens							
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;							
		Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.							
REFERENCE		1 (bases 1 to 426)							
AUTHORS		NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap .							
TITLE		National Cancer Institute, Cancer Genome Anatomy Project (CGAP),							
		Tumor Gene Index							
JOURNAL		Unpublished (1997)							
COMMENT		Contact: Robert Strausberg, Ph.D.							
		Email: cgaps-r@mail.nih.gov							
		cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo							
		, Ph.D.							
		cDNA Library Arrayed by: Greg Lennon, Ph.D.							
		DNA Sequencing by: Washington University Genome Sequencing Center							
		Clone distribution: NCI-CGAP clone distribution information can be							
		found through the I.M.A.G.E. Consortium/LLNL at:							
		www-bio.llnl.gov/bbrp/image/image.html							
		Possible reversed clone: similarity on wrong strand							
		Seq primer: -40ml3 fwd. ET from Amersham							
		High quality sequence stop: 363.							
		Location/Qualifiers							
FEATURES									

BASE COUNT		170 a	93 c	133 g	84 t		
ORIGIN							
Query Match		76.8%; Score 413.4; DB 10; Length 480;					
Best Local Similarity		97.4%; Pred. No. 6.8e-86;					
Matches 420; Conservative		0; Mismatches 11; Indels 0; Gaps 0;					
QY	108	TATCGGCCTAGACCAAGCGCTACGTAGAGCCTCTCTGAAATGATTGGCGCTATGGGGCC	167				
Db	15	TATTGGCCTAGACCAAGCGCCTATGTACAGCCTCTCTGAAAGTGAATGGCGCTATGGGGCC	74				
QY	168	GAGCAGTTCAGTGATGAAGTGGAAACAGCAACACCTGAAGAAGGGGAACCCAGCAACTCAA	227				
Db	75	GAGCAGTTCAGTGATGAAGTGGAAACAGCAACACCTGAAGAAGGGGAACCCAGCAACTCAA	134				
QY	228	CGTCAGGATCCCTGCAGCTGCTCAGGAGGGAGAGGATGAGGAGGATCTGCAGGTCAAGGG	287				
Db	135	CGTCAGGATCCCTGCAGCTGCTCAGGAGGGAGAGGATGAGGAGGATCTGCAGGTCAAGGG	194				
QY	288	CCGAAGCCTGAAGCTCATAGCCAGGAACAGGGTCACCCACAGACTGGGTGTGAGTGTGAA	347				
Db	195	CCGAAGCCTGAAGCTCATAGCCAGGAACAGGGTCACCCACAGACTGGGTGTGAGTGTGAA	254				
QY	348	GATGGTCTGTATGGGCAGGAGATGGACCCGCCAAATCCAGAGGAGGTGAAACGCCTGAA	407				
Db	255	GATGGTCTGTATGGGCAGGAGATGGACCCGCCAAATCCAGAGGAGGTGAAACGCCTGAA	314				
QY	408	GAAGGTGAAAGCAATCACAGTGTATAAGAAAGACAGCTTGAAATGATGCAGGCTGCTCC	467				
Db	315	GAAGGTGAAAGCAATCACAGTGTATAAGAAAGACAGCTTGAAATGATGCAGGCTGCTCC	374				
QY	468	TATGTTGAAATTTGTTTCATTAATAATTCCTCCATAAAGCTTTACAGCCTTCTGCAAGA	527				
Db	375	TATGTTGAAATTTGTTTCATTAATAATTCCTCCATAAAGCTTTACAGCCTTCTGCAAGA	434				
QY	528	AAAAAAAAAAAA 538					
Db	435	AAAAAAAAAAAA 445					
RESULT 15							
AA868226/c							
LOCUS							
DEFINITION							
ak48h07.s1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:1409245							
3' similar to SW:GGE2_HUMAN Q13066 GAGE-2 PROTEIN. [1] ; mRNA							
sequence.							
ACCESSION							
AA868226							
VERSION							
AA868226.1 GI:2963671							
KEYWORDS							
EST.							
SOURCE							
human.							
ORGANISM							
Homo sapiens							
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;							
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.							
1 (bases 1 to 426)							
REFERENCE							
AUTHORS							
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.							
TITLE							
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),							
Tumor Gene Index							
JOURNAL							
Unpublished (1997)							
COMMENT							
Contact: Robert Strausberg, Ph.D.							
Email: cgaps-r@mail.nih.gov							
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo							
, Ph.D.							
cDNA Library Arrayed by: Greg Lennon, Ph.D.							
DNA Sequencing by: Washington University Genome Sequencing Center							
Clone distribution: NCI-CGAP clone distribution information can be							
found through the I.M.A.G.E. Consortium/LLNL at:							
www-bio.llnl.gov/bbrp/image/image.html							
Possible reversed clone: similarity on wrong strand							
Seq primer: -40ml3 fwd. ET from Amersham							
High quality sequence stop: 363.							
Location/Qualifiers							

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 1, 2002, 06:11:45 ; Search time 201.18 seconds
(without alignments)
4591.411 Million cell updates/sec

Title: US-09-782-745-14
Perfect score: 538
Sequence: 1 ACGCAGGAGCTGTGAGGC.....CTGCAAGAAAAA 538

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_032802.*

- 1: /SIDS1/gcgdata/hold-geneseg/geneseqn-emb1/NA1980.DAT.*
- 2: /SIDS1/gcgdata/hold-geneseg/geneseqn-emb1/NA1981.DAT.*
- 3: /SIDS1/gcgdata/hold-geneseg/geneseqn-emb1/NA1982.DAT.*
- 4: /SIDS1/gcgdata/hold-geneseg/geneseqn-emb1/NA1983.DAT.*
- 5: /SIDS1/gcgdata/hold-geneseg/geneseqn-emb1/NA1984.DAT.*
- 6: /SIDS1/gcgdata/hold-geneseg/geneseqn-emb1/NA1985.DAT.*
- 7: /SIDS1/gcgdata/hold-geneseg/geneseqn-emb1/NA1986.DAT.*
- 8: /SIDS1/gcgdata/hold-geneseg/geneseqn-emb1/NA1987.DAT.*
- 9: /SIDS1/gcgdata/hold-geneseg/geneseqn-emb1/NA1988.DAT.*
- 10: /SIDS1/gcgdata/hold-geneseg/geneseqn-emb1/NA1989.DAT.*
- 11: /SIDS1/gcgdata/hold-geneseg/geneseqn-emb1/NA1990.DAT.*
- 12: /SIDS1/gcgdata/hold-geneseg/geneseqn-emb1/NA1991.DAT.*
- 13: /SIDS1/gcgdata/hold-geneseg/geneseqn-emb1/NA1992.DAT.*
- 14: /SIDS1/gcgdata/hold-geneseg/geneseqn-emb1/NA1993.DAT.*
- 15: /SIDS1/gcgdata/hold-geneseg/geneseqn-emb1/NA1994.DAT.*
- 16: /SIDS1/gcgdata/hold-geneseg/geneseqn-emb1/NA1995.DAT.*
- 17: /SIDS1/gcgdata/hold-geneseg/geneseqn-emb1/NA1996.DAT.*
- 18: /SIDS1/gcgdata/hold-geneseg/geneseqn-emb1/NA1997.DAT.*
- 19: /SIDS1/gcgdata/hold-geneseg/geneseqn-emb1/NA1998.DAT.*
- 20: /SIDS1/gcgdata/hold-geneseg/geneseqn-emb1/NA1999.DAT.*
- 21: /SIDS1/gcgdata/hold-geneseg/geneseqn-emb1/NA2000.DAT.*
- 22: /SIDS1/gcgdata/hold-geneseg/geneseqn-emb1/NA2001A.DAT.*
- 23: /SIDS1/gcgdata/hold-geneseg/geneseqn-emb1/NA2001B.DAT.*
- 24: /SIDS1/gcgdata/hold-geneseg/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	538	100.0	538	20	AA090519
2	531.6	98.8	1245	22	AA060826
3	512.8	95.3	540	20	AA090521
4	510.2	94.8	540	20	AA090523
5	504.8	93.8	532	20	AA090522
6	500.6	93.0	539	19	AA090521
7	496.8	92.3	532	19	AA090521
8	493	91.6	535	19	AA090521
9	490.6	91.2	541	19	AA090521

10	429.2	79.8	560	20	AA090520
11	421.2	78.3	560	19	AA090520
12	399.4	74.2	1024	21	AA090521
13	397.4	73.9	1024	21	AA090521
14	374.4	69.6	646	17	AA090518
15	374.4	69.6	646	20	AA090518
16	368	68.4	646	19	AA090540
17	366.4	68.1	648	16	AA090540
18	357.8	66.0	430	21	AA090540
19	301.4	56.0	530	22	AA090540
20	299.2	55.6	365	22	AA090540
21	293.6	37.8	650	21	AA090540
22	194	36.1	532	22	AA090540
23	189.2	35.2	659	22	AA090540
24	181.8	33.8	673	22	AA090540
25	173.8	32.3	611	22	AA090540
26	172.4	32.0	503	22	AA090540
27	165.6	30.8	656	21	AA090540
28	156	29.0	580	22	AA090540
29	139.8	26.0	661	23	AA090540
30	131.4	24.4	11462	19	AA090540
31	130.4	24.2	475	22	AA090540
32	127.8	23.8	665	22	AA090540
33	120.8	22.5	1051	22	AA090540
34	111.4	20.7	782	22	AA090540
35	108.4	20.1	781	22	AA090540
36	108	20.1	399	22	AA090540
37	105.4	19.6	750	23	AA090540
38	105.2	19.6	571	21	AA090540
39	101.6	18.9	515	22	AA090540
40	101.6	18.9	515	22	AA090540
41	101.6	18.9	515	22	AA090540
42	101.6	18.9	515	22	AA090540
43	100.8	18.7	214	22	AA090540
44	100.8	18.7	320	22	AA090540
45	93.4	17.4	472	21	AA090540

ALIGNMENTS

RESULT 1
ID AAX90519 standard; CDNA; 538 BP.
AC AAX90519;
DT 30-SEP-1999 (first entry)
XX GAGE-2 tumour rejection antigen clone nucleotide sequence.
DE Human leukocyte antigen; HLA-A29; tumour rejection antigen;
KW detection; therapy; pathological condition; cancer; CTL;
KW cytolytic T lymphocyte; GAGE; ss.
XX Homo sapiens.
OS
XX
PN WO9937665-A1.
XX
PD 29-JUL-1999.
XX
PF 12-JAN-1999; 99WO-US00775.
XX
PR 23-JAN-1998; 98US-0012818.
XX (LUDW-) LUDWIG INST CANCER RES.
XX Boon-Falleur T, Debacker O, Van Den Eynde B, Van Der Bruggen P;
XX cDNA encoding GAGE
XX WPI; 1999-469111/39.
PT New isolated peptides which bind to HLA-A29 molecules, which are
PT tumour rejection antigens used for detection and therapy of

PT pathological conditions, e.g. cancer
PS Example 13; Fig 4; 62pp; English.
XX
CC The present invention describes peptides which bind to human leukocyte
CC antigen (HLA) A29 (HLA-A29) molecules. The peptides are processed into
CC tumour rejection antigens. They can be used for detecting cytolytic T
CC lymphocytes (CTLs) in pathological conditions such as cancer and in
CC HLA-typing assays. Complexes of HLA-29 molecules and the peptides can
CC be used for stimulating CTLs in vivo. The present sequence represents
CC a CAGE tumour rejection antigen clone, from an example from the present
CC invention.
XX
SQ Sequence 538 BP; 160 A; 116 C; 155 G; 107 T; 0 other;

Query Match 100.0%; Score 538; DB 20; Length 538;
Best Local Similarity 100.0%; Pred. No. 3e-146;
Matches 538; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACGCCAGGAGCTGTGAGGAGCTGCTGTGCTGCTCCCTCCGACTCTTTTCTCTCT 60
DB 1 acgccaggagctgtgaggcagctgtgtgttcttcgctccgagactcttttctct 60
QY 61 ACTGAGTTCATCTGTGTAATATGAGTTGGCGAGGAAGATCGACCTATCGGCCTAGAC 120
DB 61 actgagattcatctgtgtgaatatgagttggcaggagaagatcgacctatcgccctagac 120
QY 121 CAAGACGCTACGTAGAGCTCTCGTAATGATTTGGCCCTATCGGCCCGCAGCAGTTCAGTG 180
DB 121 caagacgctacgtagagcctctgaaatgattggcctatcgcccgagcagttcagtg 180
QY 181 ATGAAGTGAACAGCAACACCTCGAAGAGGGAACACGCAACTCAACGTCAGGATCCCTG 240
DB 181 atgaagtgaaccagcaacacctgaagaagggaaccagcaactcaacgtcaggtcctg 240
QY 241 CAGCTGCTCAGAGGAGAGGATGAGGAGCATCTGAGGTCAAGGCCGCGCAGCTGAAG 300
DB 241 cagctgctcagagaggagagatgaggagcatctgcaggtcaaggggcgaagcctgaag 300
QY 301 CTCATAGCCAGGAACAGAGGTCAACCCACAGACTGGGTGTGAGTGTGAAGTGTCTCTGATG 360
DB 301 ctcatagccaggaaacagggttcacccacagactgggtgtgagctggaagatggtcctgag 360
QY 361 GGCAGGAGATGACCCGCCAATCCAGAGAGGTGAAACGCCCTGAAGAAGGTGAAAGC 420
DB 361 ggcaggagatgaccgcgcaaatccagagaggtgaaacgcctgaaagaagtgaaagac 420
QY 421 AATCACAGTGTAAAGAGACACGTTGAATGATGATGAGGCTGCTCTATGTTGGAATTT 480
DB 421 aatcacagtgttaaagaagacacgttgaaatgatgcaggtgctcctatgttggaatt 480
QY 481 TGTTCATTAATAATCTCCCAATAAAGCTTTACAGCCTTCTGCAAGAAAAA 538
DB 481 tgttcattaaaattctcccaataaagctttcacagcctctctgcgaagaaaaa 538

RESULT 2
AAS60826
ID AAS60826 standard; CDNA; 1245 BP.
AC AAS60826;
XX
XX
DT 29-JAN-2002 (first entry)
XX Human cancer agent-sensitive marker #328.
DE
XX Human; cancer cell marker; TAXOL; cytostatic; tumour; carcinoma;
KW squamous cell carcinoma; sarcoma; fibrosarcoma; leukaemia;
KW lymphocytic leukaemia; lymphoma; plasmocytoma; reticulum cell sarcoma;
KW Hodgkin's disease; glioma; ss.
XX
OS Homo sapiens.

XX WO200179556-A2.
PN XX
PD 25-OCT-2001.
XX
PF 13-APR-2001; 2001WO-US12132.
XX
PR 14-APR-2000; 2000US-197538P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Lillie J, Brown JL, Bolt A, Van Huffel C;
XX
DR WPI; 2001-602933/68.
XX
PT Novel nucleic acid, used as a marker to determine the effectiveness of
using TAXOL to treat cancer cell growth in individuals -
XX
PS Claim 1; Page 340; 527pp; English.
XX
CC The invention relates to 1046 novel nucleic acids which are used as
markers for determining the sensitivity of a cancer cell to the
anticancer agent TAXOL. Cancer cells can be treated with TAXOL when
they are shown to express one of the 242 sensitivity markers or the
cells are shown not to express one of the 804 resistance markers.
The methods can be used to determine the effectiveness of TAXOL
in the treatment of cancer cell growth in an individual. The markers
can be used as targets in developing anti-cancer agents such as
chemotherapeutic compounds. The markers can also be used as targets in
developing treatments for cancer, particularly those cancers which
display resistance to agents and exhibit expression of the markers. The
anticancer agents developed by the novel method can be used to treat
cancer. Probes based on the markers can be used to detect transcripts or
genomic sequences corresponding to the markers, in the identification of
cells or tissues which mis-express the protein. Cancers which may
be targeted include carcinoma (e.g. squamous cell carcinoma),
sarcoma (e.g. fibrosarcoma) leukaemia (e.g. lymphocytic leukaemia),
lymphoma, plasmocytoma, reticulum cell sarcoma, Hodgkin's disease and
tumours (e.g. glioma). The present sequence is one of the 1046
novel cancer cell markers.
XX
SQ Sequence 1245 BP; 382 A; 308 C; 390 G; 153 T; 12 other;

Query Match 98.8%; Score 531.6; DB 22; Length 1245;
Best Local Similarity 99.3%; Pred. No. 3.2e-144;
Matches 534; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 ACGCCAGGAGCTGTGAGGAGCTGCTGTGCTGCTCCCTCCGACTCTTTTCTCTCT 60
DB 179 acgccaggagctgtgaggcagctgtgtgttcttcgctccgagactcttttctct 238
QY 61 ACTGAGATTCATCTGTGTAATATGAGTTGGCGAGGAAGATCGACCTATCGGCCTAGAC 120
DB 239 actgagattcatctgtgtgaatatgagttggcaggagaagatcgacctattggcctagac 238
QY 121 CAAGACGCTACGTAGAGCTCTCGTAATGATTTGGCCCTATCGGCCCGCAGCAGTTCAGTG 180
DB 299 caagacgctacgtagagcctctgaaatgattgggctatcgcccgagcagttcagtg 358
QY 181 ATGAAGTGAACAGCAACACCTGAAGAAGGGGAACAGCAACTCAACGTACAGGATCCCTG 240
DB 359 atgaagtgaaccagcaacacacctgaagaagggaacacactcaacgtcaggtacctg 418
QY 241 CAGCTGCTCAGAGGAGGAGGATGAGGAGCATCTGAGGTCAAGGGCCGAGCCTGAG 300
DB 419 cagctgctcaggaggaggagaggtgaggtgaggtcaggtcaaggggccgaagcctgaag 478
QY 301 CTCATAGCCAGGAACAGGCTCAACCCACAGCTGGTGTGAGTGTGAAGTGTCTCTGATG 360
DB 479 ctcatagccaggaaacagggtcaccacacagactgggtgtgagtggaagatggtcctgag 538
QY 361 GGCAGGAGATGACCCGCCAATAATCCAGAGGAGGTGAAACGCCCTGAAGAAGGTGAAAGC 420

SO Sequence 540 BP; 159 A; 113 C; 157 G; 111 T; 0 other;

Query Match 94.8%; Score 510.2; DB 20; Length 540;
Best Local Similarity 98.0%; Pred. No. 3.6e-138;
Matches 528; Conservative 0; Mismatches 8; Indels 3; Gaps 1;

Qy 3 GCCAGGAGCTGTGAGCGAGTGTGTGTTCTGCGTCCGGACTCTTTTCTCTAC 62
Dy 1 gccagggagctgtgagcgagctgtgtgttctctgcgtccggaactcttttctctac 60
Qy 63 TGAGATTATCTGTGTGAATATAGTTGGGGAGAGATGACCC---TATCGGCTAGA 119
Dy 61 tgagattcatctgttgaaatatagttgagaggaagatgacattattattggcctaga 120
Qy 120 CCAAGACGCTACGTAGAGCTCTGCTGAATATTGGGCTATGCGGCCGAGCAGTTCACT 179
Dy 121 ccaaggcgtatgtacagcctctctgaagtgttggccctatgcgggccgagcagttccagc 180
Qy 180 GATGAAGTGGAAACCAACACCTGAAGAGAGGGGAACCAACCACTCAACGTCAGGATCCT 239
Dy 181 gatgaagtggaaacacacacacacagaggaaggggaaccagcaactcaacgtcaggtacct 240
Qy 240 GCAGCTGCTCAGGAGGAGGATGAGGAGCATTCTGCAGGTCAAGGCCCGAAGCCTGAA 299
Dy 241 gcagctgctcaggaggaggaggtgagggagcatctgcaggtcgaaggccggaagcctgaa 300
Qy 300 GCTCATAGCCAGGAACAGGCTACCCACAGACTGGGTGTGAGTGTGAAGATGGTCTGAT 359
Dy 301 gctcatagccaggaacagggtcaccacagactgggtgtgagtgatgagatggtcctgat 360
Qy 360 GGGCAGGAGATGGACCCGCAATCCAGAGAGGTGAAACGCCCTGCAAGAGAGGTGAAAG 419
Dy 361 gggcaggaggtggacccgcaaatccagagaggtgaaacgcttgaagaaagtgaaaag 420
Qy 420 CAATCACAGTGTAAAGAGACACCTTGAATGATGCAGGCTGCTCTATGTTGAAAT 479
Dy 421 caatcacaggtttaaagagacacgcttgaatgatgcaggctgctctatgttggaaat 480
Qy 480 TTGTTTCATTAAATTTCCCAATAAAGCTTTACAGCCTTCTGCAAGAAAAA 538
Dy 481 ttgttcattaaattcccaataaagctttacagccttctgcaaaaaa 539

RESULT 5

AAX90522
ID AAX90522 standard; cdNA; 532 BP.

AC AAX90522;

XX 30-SEP-1999 (first entry)

DE GAGE-5 tumour rejection antigen clone nucleotide sequence.

KW Human leukocyte antigen; HLA-A29; tumour rejection antigen;
KW detection; therapy; pathological condition; cancer; CTL;
KW cytolytic T lymphocyte; GAGE; ss.

OS Homo sapiens.

PN W09937665-A1.

XX 29-JUL-1999.

XX 12-JAN-1999; 99WO-US00775.

XX 23-JAN-1998; 98US-0012818.

XX (LUDW-) LUDWIG INST CANCER RES.

XX Boon-Falleur T., Debacker O, Van Den Eynde B, Van Der Bruggen P;
XX WPT; 1999-469111/39.

XX

PT New isolated peptides which bind to HLA-A29 molecules, which are
PT tumour rejection antigens used for detection and therapy of
PT pathological conditions, e.g. cancer
XX Example 13; Fig 4; 62pp; English.

CC The present invention describes peptides which bind to human leukocyte
CC antigen (HLA) A29 (HLA-A29) molecules. The peptides are processed into
CC tumour rejection antigens. They can be used for detecting cytolytic T
CC lymphocytes (CTLs) in pathological conditions such as cancer and in
CC HLA-typing assays. Complexes of HLA-29 molecules and the peptides can
CC be used for stimulating CTLs in vivo. The present sequence represents
CC a GAGE tumour rejection antigen clone, from an example from the present
CC invention.

XX Sequence 532 BP; 156 A; 111 C; 154 G; 111 T; 0 other;

Query Match 93.8%; Score 504.8; DB 20; Length 532;
Best Local Similarity 98.1%; Pred. No. 1.3e-136;
Matches 522; Conservative 0; Mismatches 7; Indels 3; Gaps 1;

Qy 10 ACCTGTGAGCGCAGTGTCTGTGTTCTGCGTCCGGACTCTTTTCTCTACTGAGATT 69
Dy 1 agctgtgagcgagtgctgtgtgttctgcgtccggaactctttctctactgagatt 60
Qy 70 CATCTGTGTGAATATGAGTTGGCGAGGAAGATCGACC---TATCGGCTTAGACCAAGAC 126
Dy 61 catctgtgtgaaatatagttgcgaggaagatcgacctattattggcctagaccaggc 120
Qy 127 GCTAGCTAGAGCCTCTGAAATGATTGGGCTATGCGGCCGAGCAGTTCACTGATGAAG 186
Dy 121 gctatgtacagcctctctgaagtgttggcctatggtggccgagcagttcagtgatgaag 180
Qy 187 TGGAAACCAGCAACACCTGAAGAGGGGAACACAGCAACTCAAGGTCCAGGATCCTGCAGCTG 246
Dy 181 tggaaaccagcaaacacctgaagagggaaccagcaactcaacgtcaggatcctctgcagctg 240
Qy 247 CTCAGGAGGAGAGGATGAGGGAGCATCTGCAGGTCAAGGGCCGAGCCTGAAGCTCATTA 306
Dy 241 ctcaaggaggagaggtatgagggagcatctgcagtcgaaggccgaagcctgaagctgata 300
Qy 307 GCCAGAACAGGGTCAACCCACACACTGGGTGTGAGTGTGAAGATGGTCTGTATGGCGCAGG 366
Dy 301 gccaggaacaggggtcaccacagactgggtgtgaggtggaagatggtcctgatgggcagg 360
Qy 367 AGATGGACCCGCCAAATCCAGAGGAGGTGAAACGCCCTGAAGAGGTGAAAGCAATCAC 426
Dy 361 agatggacccgccaaatcccaatccagaggaggtgaaaaacgcttgaagaggtgaaaagcaatcac 420
Qy 427 AGTGTAAAGAGACACGCTTGAATGATGCAGGCTGCTCTATGTTGGAATTTGTTCA 486
Dy 421 agtgttaaagaagcagctgtgaaatgatgcagctgctctatgttgaaaattgttca 480
Qy 487 TTAATAATCTCCCAATAAAGCTTTACAGCCTTCTGCAAGAAAAA 538
Dy 481 ttaataattctcccaataaagctttacagccttctgcaaaaaa 532

RESULT 6

AAV18721
ID AAV18721 standard; cdNA; 539 BP.

XX AAV18721;

XX 30-JUL-1998 (first entry)

XX cDNA encoding GAGE-6 tumour rejection antigen precursor.

XX GAGE tumour rejection antigen precursor; TRAP; tumour;
KW diagnosis; melanoma; antigen; cytolytic T cell clone proliferation;
KW HLA-typing assay; ss.


```
CC expression of a TRAP or presentation of a tumour rejection antigen.
CC Antigens shed into blood or urine can be observed and then used to
CC confirm a diagnosis of melanoma using cytolytic T cell clone
CC proliferation methodologies. Other uses for the processed peptides,
CC include HLA-typing assays for, e.g. skin graft or organ transplants.
XX
SQ Sequence 532 BP; 154 A; 110 C; 156 G; 112 T; 0 other;

Query Match          92.3%; Score 496.8; DB 19; Length 532;
Best Local Similarity 97.2%; Pred. No. 2.7e-134;
Matches 517; Conservative 0; Mismatches 12; Indels 3; Gaps 1;

Qy 10 AGCTGTGAGCGAGTGTGTGCTCTCGCGTCCGACTCTTTTCTCTACTGAGATT 69
Db 1 aqctgtgagcagtgctgtgttctctgcgtccgactcttttctctactagatt 60

Qy 70 CATCTGTGTGAATATGAGTTGGCGGAGAAATCGACC---TATCGGCTAGACCAAGAC 126
Db 61 catctgtgaaatatgagttggcgaggaagatcgacctattattggcctagaccaagc 120

Qy 127 GCTAGCTAGAGCTCTCTGAATGATTGGCTATGCGGCCGAGCAGTTCAGTGTGAAG 186
Db 121 ggtatgtacagcctctctgaagtgtattggcctatgcgcccagcagttcagtgatgaag 180

Qy 187 TGGAAACCAACACCTGAAAGAGGGGAACCACTCAACGTCAGGATCCTGCAGCTG 246
Db 181 tggaaacagagcctctgaagaaggggaaacacgacaaactcaacgtcaggatcctgcagctg 240

Qy 247 CTCAGGAGGAGGATGAGGGAGCATCTGCAGGTCAAGGCCGAGCCTGAAGTTCATA 306
Db 241 ctcadgaggaggagatgagggagcatctgcagtgcaaggccgagcctgaagctgata 300

Qy 307 GCAGGAACAGGGTCAACCCACAGACTGGGTGTGAGTGTGAAGATGGTCTGTATGGCAGG 366
Db 301 gccaggaaacagggctcaacacagactgggtgtgagtggaagatggtcctgatggcaggg 360

Qy 367 AGATGACCCGCCAAATCCAGAGAGGTGAAACGCCTGAGAAAGGTGAAAGCAATCAC 426
Db 361 agatgaccgcgcgaataccagaggaggtgaaacgcctgaagaaggtgaaagcaatcac 420

Qy 427 AGTGTAAAGAGACACGTTGAAATGATGCAGCTGCTCTATCTGTGAAATTTCTTCA 486
Db 421 agtgttaaagaggcagcttgaaatgatgcagcgtctctcatgttggaaatttgcttca 480

Qy 487 TTAATAATCTCCCAATAAAGCTTTACAGCCTTCTGCAAGAAAAAATAAAAAA 538
Db 481 ttaaaattctcccaataaagctttacagcctctctgcaagaaaaaataaaaaa 532

RESULT 8
AAV18717
ID AAV18717 standard; cDNA; 535 BP.
XX
AC AAV18717;
XX
DT 30-JUL-1998 (first entry)
XX
DE cDNA encoding GAGE-2 tumour rejection antigen precursor.
XX
KW GAGE tumour rejection antigen precursor; TRAP; tumour;
KW diagnosis; melanoma; antigen; cytolytic T cell clone proliferation;
KW HLA-typing assay; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 81..431
FT /tag= a
FT /transl_except= (pos:192..194, aa:Ala)
FT /transl_except= (pos:195..197, aa:Thr)
XX
PN WO9749417-Al.
```

```
XX 31-DEC-1997.
PD
XX 23-JUN-1997; 97WO-US10850.
PF
XX 24-JUN-1996; 96US-0669161.
PR
XX (LUDW-) LUDWIG INST CANCER RES.
PA
XX Boon-Falleur T, Debacker O, Van Den Eynde B;
XX WPI; 1998-076905/07.
XX P-PSDB; AAW47599.
DR
XX Isolated nucleic acid encoding GAGE tumour rejection antigen
XX precursor - processed by HLA-Cw6 molecules into peptides, useful to
XX diagnose melanomas
XX
XX Example 13; Fig 4; 60pp; English.
XX
XX The present sequence encodes a GAGE-2 tumour rejection antigen
XX precursor (TRAP). The protein is expressed in a number of tumours. In
XX contrast the only normal tissue which expresses GAGE TRAP protein is
XX testis. Several GAGE TRAPs have been identified (see AAV18717-21). The
XX major difference between these proteins and GAGE-1 (AAV05540) is the
XX absence of a stretch of 143 bases located at position 379 to 521 of the
XX GAGE-1 TRAP sequence. The rest of the sequences show mismatches at
XX various position, with the exception of GAGE-3 whose 5' end is totally
XX different from the other GAGE cDNAs for the first 112 bases. This
XX region of GAGE-3 cDNA contains a long repeat and a hairpin structure.
XX The antigens can be used to diagnose melanomas, characterised by
XX expression of a TRAP or presentation of a tumour rejection antigen.
XX Antigens shed into blood or urine can be observed and then used to
XX confirm a diagnosis of melanoma using cytolytic T cell clone
XX proliferation methodologies. Other uses for the processed peptides,
XX include HLA-typing assays for, e.g. skin graft or organ transplants.
XX
XX Sequence 535 BP; 158 A; 112 C; 156 G; 109 T; 0 other;
SQ

Query Match          91.6%; Score 493; DB 19; Length 535;
Best Local Similarity 97.2%; Pred. No. 3.5e-133;
Matches 524; Conservative 0; Mismatches 10; Indels 5; Gaps 2;

Qy 1 AGCCAGGAGCTGTGAGGAGCTGCTGTGTGCTCTGCCGTCCGACTCTTTTCTCTC- 59
Db 1 acgccaggagctgtgaggcagtgctgtg----ttcttgcgtccgagaccttttctct 56

Qy 60 TACTGAGATTCACTGTGTGAAATATGAGTTGGCGGAGGAAGATCGACCTATCGCCCTAGA 119
Db 57 tactgagattcatctgtgtgaaatatgagttggcgaggaagatcgacctatcgccctaga 116

Qy 120 CCAAGACGCTACGTAGAGCCTCTCTGAATGATTGGGCTCTATCGGCCCGAGAGTTCAGT 179
Db 117 ccaagacgctacgtagagcctctgaaatgattggcctatggtcgccgagcagttcagt 176

Qy 180 GATGAAGTGAACCAACCACTGAAGAGGGGAACCAAGCAACTCAAGCTCAGGATCCT 239
Db 177 gatgaagtgaaccagagcctctgaaagaagggaaccagcaactcaacgtcaggatcct 236

Qy 240 GCAGTGTCTCAGGAGGAGAGGATGAGGAGCATCTGCGAGTCAAGGGCCGAAGCCCTGAA 299
Db 237 gcagtcgctcaggaggaggaggtgagggagcatctgcaggtcaaaagggccgaagcctgaa 296

Qy 300 GCTCATAGCCAGGAACAGGGTCAACCCACAGACTGGGTGTGAGTGTGAAGATGGTCTCAT 359
Db 297 gctgagagccaggaacagggtcacccacagactgggtgtgagtgtaagatggtcctgat 356

Qy 360 GGGCAGGAGATGAGCCCGCAATCCAGAGAGGTTGAACCGCTCAAGAGAGTGAAGAG 419
Db 357 gggcagagatgagcccgcaaatccagagaggtgaaacacgctcctgaaagtgaaaag 416

Qy 420 CAATCACAGTGTATAAAGAGACACGTTGAAATGATGCAGGCTGCTCTCTATGTGGAAT 479
```

Db 417 caatcacagtgttaaaagagacatgttgaaatgatgcaggctgctcctcatgttggaaat 476
Qy 480 TTGTTTCATTAATAATTCCTCCCAATAAAGCTTTACAGCCTTCTGCAAGCAAAAAA 538
Db 477 ttgttcattaaattctcccaataaagctttacagccttctgcgaagaaaaa 535

RESULT 9
AAV18719
ID AAV18719 standard; cDNA; 541 BP.
AC AAV18719;
XX
XX
30-JUL-1998 (first entry)
XX
XX
cDNA encoding GAGE-4 tumour rejection antigen precursor.
XX
KW GAGE tumour rejection antigen precursor; TRAP; tumour;
KW diagnosis; melanoma; antigen; cytolytic T cell clone proliferation;
KW HLA-typing assay; ss.

OS Homo sapiens.
FH Key Location/Qualifiers
FT CDS 83..435
FT /*tag= a
FT /transl_except= (pos:197..199, aa:Ala)
FT /transl_except= (pos:200..202, aa:Thr)

XX WO9749417-A1.
XX
XX 31-DEC-1997. 97WO-US10850.
XX
XX 23-JUN-1997; 97WO-US10850.
XX
XX 24-JUN-1996; 96US-0669161.
XX
XX (LUDW-) LUDWIG INST CANCER RES.
XX
XX Boon-Falleur T, Debacker O, Van Den Eynde B;
XX
XX WPI; 1998-076905/07.
XX P-PSDB; AAW47601.

XX Isolated nucleic acid encoding GAGE tumour rejection antigen
XX precursor - processed by HLA-Cw6 molecules into peptides, useful to
XX diagnose melanomas
XX
XX Example 13; Fig 4; 60pp; English.

XX The present sequence encodes a GAGE-4 tumour rejection antigen
XX precursor (TRAP). The protein is expressed in a number of tumours. In
XX contrast the only normal tissue which expresses GAGE TRAP protein is
XX testis. Several GAGE TRAPs have been identified (see AAV18717-21). The
XX major difference between these proteins and GAGE-1 (AAV05540) is the
XX absence of a stretch of 143 bases located at position 379 to 521 of the
XX GAGE-1 TRAP sequence. The rest of the sequences show mismatches at
XX various position, with the exception of GAGE-3 whose 5' end is totally
XX different from the other GAGE cDNAs for the first 112 bases. This
XX region of GAGE-3 cDNA contains a long repeat and a hairpin structure.
XX The antigens can be used to diagnose melanomas, characterised by
XX expression of a TRAP or presentation of a tumour rejection antigen.
XX Antigens shed into blood or urine can be observed and then used to
XX confirm a diagnosis of melanoma using cytolytic T cell clone
XX proliferation methodologies. Other uses for the processed peptides,
XX include HLA-typing assays for, e.g. skin graft or organ transplants.
XX
XX Sequence 541 BP; 158 A; 117 C; 154 G; 112 T; 0 other;

Query Match 91.2%; Score 490.6; DB 19; Length 541;
Best Local Similarity 96.7%; Pred. No. 1.8e-132;

Matches 523; Conservative 0; Mismatches 14; Indels 4; Gaps 2;
Qy 2 CGCCAGGAGCTGTGAGCAGTCTGTGTGTGTTCTCTGCCCTCGGACTCTTTTCTCTCTA 61
Db 1 cgcacggagctgtgaggcagtgctgtgtggttcctgcgcgtcccaactcttttccctcta 60
Qy 62 CTGAGATTCACTGTGTGAAATATGAGTTGGCCAGGAGNAGATCGACC---TATCGCCCTAG 118
Db 61 ctgagattcatctgttgaaatgatgagttggcgaggaagatcgacctattattggccctag 120
Qy 119 ACCAAGACGCTACGTAGAGCCTCTCAAAATGATTGGGCCCTATGGCGCCGAGCAGTTCAG 178
Db 121 accaaagcgctatgtacagcctcctctgaaatgatgggacctatgcggcccgagcagttcag 180
Qy 179 TGATGAAGTGGAAACCAACCAACCTGAAGAAGGGGAACCAACCACTCAACCTCAGGATCC 238
Db 181 tgaatgaagtggaaaccagagcctctctgaagaaggggaaaccagcaactcaacgtcaggatcc 240
Qy 239 TGCAGCTGCTCAGGAGGAGAGATGAGGAGGCATCTGCAGGTCAAGGCCGAGCCCTGA 298
Db 241 tgcagctgctcagggaggagagatgagggagcatctgcaggtcaaggccgaagcctga 300
Qy 299 AGCTCATAGCCAGGAACAGGGTCCACCACAGACTGGGTGTGAGTGTGAAGATGGTCTCTGA 358
Db 301 agctgatagccagggaacagggtccaccacagactggtgtgagtgtagagatggtcctga 360
Qy 359 TGGCAGGAGATGGACCCGCCCAATCCAGAGAGGTGAAACCGCTCAAGAGGTGAAAA 418
Db 361 tgggcaggagatggacccgccaaatccagagagggtgaaaaagcgtgaagaaggtgaaaa 420
Qy 419 GCAATCACAGTGTCTAAAGAAGACAGTTG-AAATGATGCAGGCTGCTCTATCTTGAA 477
Db 421 gcaatcacagtggttaaaagagagcaggttgaaaatgatgcaggtctgcccatgtctggaa 480
Qy 478 ATTTGTTTCAATAAATTCCTCCCAATAAAGCTTTACAGCCTTCTCAAGAAAAA 537
Db 481 atttgttcattaaattctcccaataaagcttttacagcctctctgcaaaaaa 540
Qy 538 A 538
Db 541 a 541

RESULT 10
AAV90520
ID AAV90520 standard; cDNA; 560 BP.
XX
XX AC AAV90520;
XX
XX DT 30-SEP-1999 (first entry)
XX

XX GAGE-3 tumour rejection antigen clone nucleotide sequence.
XX
XX DE
XX DE Human leukocyte antigen; HLA-A29; tumour rejection antigen;
XX KW detection; therapy; pathological condition; cancer; CTL;
XX KW cytolytic T lymphocyte; GAGE; ss.
XX
XX OS Homo sapiens.

XX
XX
XX PN WO9937665-A1.
XX
XX PD 29-JUL-1999.
XX
XX
XX PF 12-JAN-1999; 99WO-US00775.
XX
XX PR 23-JAN-1998; 98US-0012818.
XX
XX PA (LUDW-) LUDWIG INST CANCER RES.

XX
XX PI Boon-Falleur T, Debacker O, Van Den Eynde B, Van Der Bruggen P;
XX
XX DR WPI; 1999-469111/39.
XX

PT	New isolated peptides which bind to HLA-A29 molecules, which are
PT	tumour rejection antigens used for detection and therapy of
PT	pathological conditions, e.g. cancer
XX	
PS	Example 13; Fig 4; 62pp; English.
XX	
CC	The present invention describes peptides which bind to human leukocyte
CC	antigen (HLA) A29 (HLA-A29) molecules. The peptides are processed into
CC	tumour rejection antigens. They can be used for detecting cytolytic T
CC	lymphocytes (CTLs) in pathological conditions such as cancer and in
CC	HLA-typing assays. Complexes of HLA-29 molecules and the peptides can
CC	be used for stimulating CTLs in vivo. The present sequence represents
CC	a GAGE tumour rejection antigen clone, from an example from the present
CC	invention.
XX	
SQ	Sequence 560 BP; 177 A; 119 C; 150 G; 114 T; 0 other;

[illegible]

RESULT	11
AAV18718	
ID	AAV18718 standard; cDNA; 560 BP.
XX	
XX	
AAV18718;	
AC	
XX	
XX	
DT	
XX	
XX	
DE	30-JUL-1998 (first entry)
XX	
XX	
XX	CDNA encoding GAGE-3 tumour rejection antigen precursor.
KW	
KW	GAGE tumour rejection antigen precursor; TRAP; tumour;
KW	diagnosis; melanoma; antigen; cytolytic T cell clone proliferation;
KW	HLA-typing assay; ss.
XX	
XX	

OS	Homo sapiens.
XX	
XX	Key
XX	Location/Qualifiers
FT	100..457
FT	/*tag= a
FT	/transl_except= (pos:217..219, aa:Ala)
FT	/transl_except= (pos:220..222, aa:Thr)
XX	
XX	WO9749417-A1.
PN	
XX	
XX	31-DEC-1997.
PD	
XX	
XX	23-JUN-1997; 97WO-US10850.
PF	
XX	
XX	24-JUN-1996; 96US-0669161.
PR	
XX	
XX	(LUDW-) LUDWIG`INST CANCER RES.
PA	
XX	
XX	Boon-Falleur T, Debacker O, Van Den Eynde B;
PI	
XX	
XX	WPI; 1998-076905/07.
DR	
XX	P-PSDB; AAW47600.
DR	
XX	
XX	Isolated nucleic acid encoding GAGE tumour rejection antigen
PT	precursor - processed by HLA-Cw6 molecules into peptides, useful to
PT	diagnose melanomas
PT	
XX	
XX	Example 13; Fig 4; 60pp; English.
PS	
XX	
CC	The present sequence encodes a GAGE-3 tumour tumour rejection antigen
CC	precursor (TRAP). The protein is expressed in a number of tumours. In
CC	contrast the only normal tissue which expresses GAGE TRAP protein is
CC	testis. Several GAGE TRAPS have been identified (see AAV18717-21). The
CC	major difference between these proteins and GAGE-1 (AAV05540) is the
CC	absence of a stretch of 143 bases located at position 379 to 521 of the
CC	GAGE-1 TRAP sequence. The rest of the sequences show mismatches at
CC	various position, with the exception of GAGE-3 whose 5' end is totally
CC	different from the other GAGE cDNAs for the first 112 bases. This
CC	region of GAGE-3 cDNA contains a long repeat and a hairpin structure.
CC	The antigens can be used to diagnose melanomas, characterised by
CC	expression of a TRAP or presentation of a tumour rejection antigen.
CC	Antigens shed into blood or urine can be observed and then used to
CC	confirm a diagnosis of melanoma using cytolytic T cell clone
CC	proliferation methodologies. Other uses for the processed peptides,
CC	include HLA-typing assays for, e.g. skin graft or organ transplants.
CC	
XX	
XX	Sequence 560 BP; 175 A; 119 C; 152 G; 114 T; 0 other;
SQ	

	Query Match	78.3%	Score 421.2;	DB 19;	Length 560;
	Best Local Similarity	91.2%;	Pred. No. 2.6e-112;		
	Matches 447;	Conservative	0;	Mismatches 43;	Indels 0;
QY	49	TCTTTTCTCTCTACTGAGATTCTACTGTGTAATATGATGTGGCGAGGAGATCGACCT	108		
Db	71	tgtctctcgccaactcatatttcacacagatgaatctcagtadaggaataatcgacctatt	130		
QY	109	ATCGGCCTAGACCAAGACGCTACGTAGAGCCTCTCTGAATGATTGGGCCCTATCGGGCCCG	168		
Db	131	attggcctagacaaagcgctatgcagcctctggaagtattggcctatcgggccc	190		
QY	169	AGCAGTTCAGTGATGAAGTGGAGCCAGCAACACCTTGAAGAAGGGGAACAGCAACTCAAC	228		
Db	191	agcagttcagtgatgaagtggaaaccagagcctctctgaagaaaggggaaccagaaactcaac	250		
QY	229	GTCAAGATCCTCGAGCTGCTCAGGAGGGAGAGGATGAGGAGGCATCTGCAAGGTCAAGGGC	288		
Db	251	gtcaggatcctgcagctgctcaggagggagagatgagggagcatctgcaggtcgaagggc	310		
QY	289	CGAAGCCTGAAGGCTATAGCCAGAAACAGGGTCACCACAGACTGGGTGTGAGTGTGAAG	348		
Db	311	cgaagcctgaagctgatagcagaaacaggttcaccacagactgggtgagtgtgaaag	370		

QY 349 ATGGTCTGATGGCAGGAGATGGACCCGCCAAATCAGAGGAGTGAACCGCTGAAG 408
DB 371 atggtctctgagggcaggagatgagcccgccaaatccagaggagtgaacacgctgaag 430
QY 409 AAGGTGAAGAAGCAATCACAGTGTAAAGAAAGACACACCTTGAATGATGACGGCTGCTCT 468
DB 431 aagggtgaagaagcaatcacagtgtaaagaagggcagcgttgaaatgatgcaggctgctct 490
QY 469 ATGTTGGAATTTGTTTCATTAATCTCCCAATTAAGCTTTACAGCCTTCTGCAAGAA 528
DB 491 atggtggaatattgtttcatcaaatctcccaataaaggctttacagccttctgcaagaa 550
QY 529 AAAAAA 538
DB 551 aaaaaaaaaa 560

RESULT 12
AAZ97217
ID AAZ97217 standard; cDNA; 1024 BP.
XX
AC AAZ97217;
DT 18-APR-2000 (first entry)
DE Human prostate cancer differentially expressed gene #78.
XX
KW Prostate cancer specific gene; cancer; tumour progression; diagnose;
KW hyperproliferative cell growth; prostatic disorder; treatment;
KW metastatic prostate cancer; benign prostate hyperplasia; BPH; ss.
XX
OS Homo sapiens.
XX
PN WO9964594-A2.
XX
PD 16-DEC-1999.
XX
PF 10-JUN-1999; 99WO-US13181.
XX
PR 11-JUN-1998; 98US-0088877.
PR 09-JUN-1999; 99US-0088877.
XX
XX (CHIR) CHIRON CORP.
XX
PI Astel JH, Carroll E, Endege WO, Ford DM, Monahan JE, Schlegel R;
PI Steinmann KE, Zhang J;
XX
XX WPI; 2000-116541/10.
XX
XX New isolated prostate cancer specific nucleic acids, used to develop
PT products for the diagnosis and treatment of cancer -
XX
XX Claim 2; Page 107; 212pp; English.
XX
XX This sequence represents a prostate cancer specific nucleic acid
CC sequence. The invention relates to a method for diagnosing cancer,
CC tumour progression, hyperproliferative cell growth or accompanying
CC biological and physical manifestations. The method involves contacting
CC the biological sample with a probe that comprises a sequence capable of
CC hybridising to any of the 339 nucleotide sequences given in the
CC specification (see AAZ97140-297478) and detecting duplex formation. The
CC products and methods of the invention can be used for the diagnosis,
CC prognosis, and treatment of cancer, tumour progression,
CC hyperproliferative cell growth, and accompanying physical and biological
CC manifestations. They can be used particularly for prostatic disorders
CC such as metastatic prostate cancer, localised prostate cancer, or benign
CC prostate hyperplasia (BPH).
XX
SQ Sequence 1024 BP; 272 A; 243 C; 275 G; 207 T; 27 other;

Query Match 74.28; Score 399.4; DB 21; Length 1024;
Best Local Similarity 98.58; Pred. No. 7.3e-106;

Matches 403; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 130 ACGTAGAGCTCTCCCTGAAATGATGGGCCCTATCGCGCCCGCAGAGTTTCAGTGAAGCTGG 189
DB 100 aggtacagcctccctgaaatgattgggcccctatgcccgcagagcagttcagtgatgaagcgg 159
QY 190 AACGAGCAACACCTCAAGAGGGGAAACGAGCAACTCAACGTCAAGATCCTGACGCTGCTC 249
DB 160 aaccagcaacacctgaagaagggaaccagcaactcaacytcaggatcctgcagctgctc 219
QY 250 AGGAGGAGAGGATGAGGGAGCATCTGCAGGTCAAGGCGCGAGCCTGAAGCTCATAGCC 309
DB 220 agggaggagagatgagggagcatctgcaggtcgaagggcccgaagcctgaagcccatagcc 279
QY 310 AGGAACAGGGTCACCCACAGACTGGGTGTGAGTGTGAAGATGCTCTGATGGCAGAGAGA 369
DB 280 aggaacagggtcacccacagactgggtgagtgagtgagtgagtgagtgagtgagtgagtgag 339
QY 370 TGGACCCGCCAAATCCAGAGAGGTGAACCGCTGAAGAGTGAACCAATTCACAGT 429
DB 340 tggacccgcgaatccagaggaggtgaacacgcctgaagaagtgaaaagcaatccacagt 399
QY 430 GTTAAAGAAGACACGCTTGAATGATGCAGGCTGCTCTATGTTGGAATTTGTTCAATTA 489
DB 400 gttaaagaagaagcagcttgaaaatgatgcaggctgctcctatgttgaaaattgttcatta 459
QY 490 AAATTTCTCCCAATAAGCTTTACAGCCTTCTGCAAGAAAGAAAAA 538
DB 460 aaattctcccaataaagctttacagccttctgtaaaaaaaaaaaaaa 508

RESULT 13
AAZ97216/c
ID AAZ97216 standard; cDNA; 1024 BP.
XX
AC AAZ97216;
XX
DT 18-APR-2000 (first entry)
DE Human prostate cancer differentially expressed gene #77.
XX
KW Prostate cancer specific gene; cancer; tumour progression; diagnose;
KW hyperproliferative cell growth; prostatic disorder; treatment;
KW metastatic prostate cancer; benign prostate hyperplasia; BPH; ss.
XX
OS Homo sapiens.
XX
PN WO9964594-A2.
XX
PD 16-DEC-1999.
XX
PF 10-JUN-1999; 99WO-US13181.
XX
PR 11-JUN-1998; 98US-0088877.
PR 09-JUN-1999; 99US-0088877.
XX
XX (CHIR) CHIRON CORP.
XX
PI Astel JH, Carroll E, Endege WO, Ford DM, Monahan JE, Schlegel R;
PI Steinmann KE, Zhang J;
XX
XX WPI; 2000-116541/10.
XX
XX New isolated prostate cancer specific nucleic acids, used to develop
PT products for the diagnosis and treatment of cancer -
XX
XX Claim 2; Page 107; 212pp; English.
XX
XX This sequence represents a prostate cancer specific nucleic acid
CC sequence. The invention relates to a method for diagnosing cancer,
CC tumour progression, hyperproliferative cell growth or accompanying
CC biological and physical manifestations. The method involves contacting
CC the biological sample with a probe that comprises a sequence capable of
CC hybridising to any of the 339 nucleotide sequences given in the
CC specification (see AAZ97140-297478) and detecting duplex formation. The
CC products and methods of the invention can be used for the diagnosis,
CC prognosis, and treatment of cancer, tumour progression,
CC hyperproliferative cell growth, and accompanying physical and biological
CC manifestations. They can be used particularly for prostatic disorders
CC such as metastatic prostate cancer, localised prostate cancer, or benign
CC prostate hyperplasia (BPH).
XX
SQ Sequence 1024 BP; 272 A; 243 C; 275 G; 207 T; 27 other;

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 1, 2002, 06:29:18 ; Search time 14.38 Seconds
(without alignments)
775.130 Million cell updates/sec

Title: US-09-782-745-27
Perfect score: 646
Sequence: 1 MSWGRSTYRPRRYVEPP.....DPNPVEVKTPEEGEKQSQ 116
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues
Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR71:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	119	18.4	1300	2 T03166	probable immediate
2	110	17.0	139	2 S70010	glutamate/proline-
3	110	17.0	221	2 S70009	glutamate/proline-
4	106	16.4	301	2 E29149	proline-rich prote
5	105.5	16.3	260	2 S22373	proline-rich prote
6	104	16.1	396	1 A58938	surface protein rh
7	103.5	16.0	171	2 A27307	proline-rich phosph
8	102.5	15.9	815	2 B30843	glutenin high mole
9	102.5	15.9	815	2 JN0689	glutenin, high-mol
10	102.5	15.9	830	2 S15720	glutenin high mole
11	102.5	15.9	848	2 S02262	glutenin high mole
12	102	15.8	652	2 PC4014	calcium binding 14
13	102	15.8	1560	2 T30282	calcium-binding pr
14	101.5	15.7	300	2 S19560	proline-rich prote
15	101	15.6	754	2 B97586	hypothetical prote
16	101	15.6	754	2 AC2807	OmpA family protei
17	100.5	15.6	605	1 FWSVBA	beta-conglycinin a
18	100.5	15.6	847	2 F96531	hypothetical prote
19	100	15.5	1881	2 H95076	zinc metalloprotei
20	99.5	15.4	309	2 S10889	proline-rich prote
21	99.5	15.4	548	1 QFPGI	neurofilament trip
22	99.5	15.4	605	2 S20007	beta-conglycinin a
23	99.5	15.4	1110	2 T51116	NF-180 - sea lampr
24	98	15.2	212	2 B36298	proline-rich prote
25	98	15.2	838	1 EEWTHW	glutenin, high mol
26	98	15.2	1606	2 T34073	paranenin - chicke
27	98	15.2	2649	2 T51023	hypothetical prote
28	97.5	15.1	166	2 B25372	salivary proline-r
29	97.5	15.1	3436	2 S55659	tegument protein 6

30	97	15.0	204	2 A39066	proline-rich prote
31	96.5	14.9	229	1 W4WL8	E4 protein - human
32	96.5	14.9	617	2 S42719	actin-binding prot
33	96	14.9	543	1 QFMSL	neurofilament trip
34	96	14.9	554	2 JW0094	neurofilament prot
35	96	14.9	1876	2 E97944	zinc metalloprotei
36	95.5	14.8	544	2 S07144	neurofilament trip
37	95.5	14.8	634	1 GBRTX1	matrix glycoprotei
38	94.5	14.6	220	2 A36298	proline-rich prote
39	94.5	14.6	530	2 T48627	hypothetical prote
40	94	14.6	183	2 T20980	hypothetical prote
41	94	14.6	1263	2 T13805	spalt-related prot
42	93.5	14.5	245	1 W4WL5	E4 protein - human
43	93.5	14.5	310	1 PHUSD	salivary proline-r
44	93.5	14.5	325	2 T16324	hypothetical prote
45	93	14.4	206	1 PIRT3	acidic proline-ric

ALIGNMENTS

RESULT 1
T03166
probable immediate early protein - alcelaphine herpesvirus 1
C:Species: alcelaphine herpesvirus 1
C:Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 24-Nov-1999
C:Accession: T03166
R:Ensser, A.; Pflanz, R.; Fleckenstein, B.
J. Virol. 71, 6517-6525, 1997
A:Title: Primary structure of the alcelaphine herpesvirus 1 genome.
A:Reference number: 214840; MUID:97404659
A:Accession: T03166
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-1300 <ENS>
A:Cross-references: EMBL:AF005370; NID:g2337967; PIDN:AAC58118.1; PID:g2338034
C:Superfamily: collagen alpha 2(I) chain; fibrillar collagen carboxyl-terminal hom

Query Match 18.4%; Score 119; DB 2; Length 1300;
Best Local Similarity 35.2%; Pred. No. 0.18;
Matches 45; Conservative 7; Mismatches 44; Indels 32; Gaps 7;

QY	17	VEPEMIGMPREQSDEVEPATPE--EGE-----PATQRDDPAAQEG-----E 59
DB	527	LEGPEGEGPEGPEGE--GPEGPEGEGPEGPEGRDSDPGCAQEGPEGPEE 585
QY	60	DEGASAGGKPKPEAHSQEQHPQTGCECDG-----PDQEQMDPPN-----PEEVKT 106
DB	586	DEGPEGPEGPEGPEGPEGPEGPEGPEGPEGPEGPEGPEGPEGPEGPEGPKG 645
QY	107	PEEGEKQS 114
DB	646	P-EGECQS 652

RESULT 2
S70010
glutamate/proline-rich protein (clone BB.1.4.1) - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 06-Dec-1996 #sequence_revision 13-Mar-1997 #text_change 05-Nov-1999
C:Accession: S70010
R:Geertman, R.; McMahon, A.; Sabban, E.L.
Biochim. Biophys. Acta 1306, 147-152, 1996
A:Title: Cloning and characterization of cDNAs for novel proteins with glutamic acid
A:Reference number: S70009; MUID:96221285
A:Accession: S70010
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-139 <GEE>
A:Cross-references: EMBL:U40628; NID:g1184695; PIDN:AAB05668.1; PID:g1184696

Db 428 GQGQGPQGCGEQGPQGQGGQGPQGQPGYPTSPQSGQGGQPGYPTSPQSGQLQ 487

Qy 107 PEEGEKQSQ 115

Db 488 PAQGQPGQ 496

RESULT 11

S02262
glutenin high molecular weight chain Dx5 - wheat
C:Species: Triticum aestivum (common wheat)
C:Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 03-Feb-1994
C:Accession: S02262
R:Anderson, O.D.; Greene, F.C.; Yip, R.E.; Halford, N.G.; Shewry, P.R.; Malpica-Romero,
Nucleic Acids Res. 17, 461-462, 1989
A:Title: Nucleotide sequences of the two high-molecular-weight glutenin genes from the
A:Reference number: S02262; MUID:89098419
A:Accession: S02262
A:Molecule type: DNA
A:Residues: 1-848 <AND>
A:Cross-references: EMBL:X12928
C:Genetics:
A:Gene: Glu-D1-1b
C:Superfamily: glutenin

Query Match 15.9%; Score 102.5; DB 2; Length 848;
Best Local Similarity 23.4%; Pred. No. 2;
Matches 33; Conservative 24; Mismatches 53; Indels 31; Gaps 6;

Qy 4 RGRSTYRP-----RPRRYVEPPE-----MIGMRPEQFS---DEVFATPEEGE 44

Db 491 QGQPGYPTSPQSGQGGQGPQGGQGPQGGYPTSPQPGQGGQGDYDPTSPQPGQGGQ 550

Qy 45 PATQRDPAQAAGE-----DEGASAGQGPKE--AHSGQGHGHPGTGCECDGPDGQ 94

Db 551 QFGQLQPAQGGQGGQLAQGGQGGQPAQVQGGQRPAGQGGQGGQGGQLGGQGGQ 610

Qy 95 EMDPNPEEVKTPPEGEKQSQ 115

Db 611 Q--PGQGGQGGQPAQGGQGGQ 629

RESULT 12

PC4014
calcium binding 140k protein - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 21-Jul-1995 #sequence_revision 31-Mar-1997 #text_change 05-Nov-1999
C:Accession: PC4014
R:Naved, A.F.; Ozawa, M.; Yu, S.; Miyauchi, T.; Muramatsu, H.; Muramatsu, T.
Cell Struct. Funct. 20, 133-141, 1995
A:Title: CBP-140, a novel endoplasmic reticulum resident Ca2+-binding protein with a car
A:Reference number: PC4014; MUID:95368740
A:Accession: PC4014
A:Molecule type: mRNA
A:Residues: 1-652 <NAY>
A:Cross-references: GB:S78797; NID:g1041986; PIDN:AAB35051.1; PID:g1041987
C:Comment: The sequence 1-80 has high homology to that of 70k heat shock protein.
C:Keywords: calcium binding; endoplasmic reticulum; heat shock; stress-induced protein
F:649-652/Region: endoplasmic reticulum retention signal #status atypical

Query Match 15.8%; Score 102; DB 2; Length 652;
Best Local Similarity 34.1%; Pred. No. 1;
Matches 31; Conservative 8; Mismatches 36; Indels 16; Gaps 4;

Qy 25 EMRPQFSDEVFATPEEGEPATQRDPAQAAGE--DEGASAGQGPKEAHSQGGHPQ 82

Db 270 PAEQGELKEEAPPAETSQPPSPKGDAAAREGKPKDEKES---GDKPEAQPNKKG---- 323

Qy 83 TGCECDGPDGQEMDPNPEEVKTPPEGEKQ 113

Db 324 ----QAPEGA---APAEEDKKLPARKQ 346

RESULT 13

T30282
calcium-binding protein - sea urchin (Strongylocentrotus purpuratus)
C:Species: Strongylocentrotus purpuratus (purple urchin)
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 11-May-2000
C:Accession: T30282
R:Soltyisk-Espanola, M.; Klinzing, D.C.; Pfarr, K.; Burke, R.D.; Ernst, S.G.
Dev. Biol. 165, 73-85, 1994
A:Title: Endo16, a large multidomain protein found on the surface and ECM of en
A:Reference number: Z20805; MUID:94374583
A:Accession: T30282
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1560 <SOI>
A:Cross-references: EMBL:L34680; NID:g511893; PID:g511894; PIDN:AAA30047.1
C:Genetics:
A:Note: endo16

Query Match 15.8%; Score 102; DB 2; Length 1560;
Best Local Similarity 32.0%; Pred. No. 3.8;
Matches 32; Conservative 13; Mismatches 35; Indels 20; Gaps 5;

Qy 34 EYEPATPEEGEPATQRDPAQAAGEDEGASAGQGPKEAHSQEGHHPQTGCECE----- 88

Db 1362 EVKTSNOEGSEVQEESQDPQEMEGEGQG--SESEGEPEEPGEQG--PQEQVSEGEENE 1419

Qy 89 -----DGP-----DQEMDPNPE-EVKTPPEGEKQSQ 115

Db 1420 SEEEVEEPQGVMESEGGQSESEPEPEGEPESEGEQE 1459

RESULT 14

SL9560
proline-rich protein MP4 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 20-Aug-1999
C:Accession: SL9560; S22570
R:Roberts, S.G.E.; Layfield, R.; Bannister, A.J.; McDonald, C.J.
Eur. J. Biochem. 202, 969-974, 1991
A:Title: Gene sequence of mouse B-type proline-rich protein MP4. Transcriptional
A:Reference number: SL9560; MUID:92111548
A:Accession: SL9560
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-300 <ROB>
A:Cross-references: GB:X58438; NID:g53181; PIDN:CAA41344.1; PID:g53182
R:Roberts, S.G.E.; Layfield, R.; McDonald, C.J.
Nucleic Acids Res. 19, 5205-5211, 1991
A:Title: The mouse proline-rich protein MP6 promoter binds isoprenaline-inducibl
A:Reference number: S22570; MUID:92020206
A:Accession: S22570
A:Molecule type: DNA
A:Residues: 1-14 <RO>
A:Cross-references: EMBL:X61126
C:Superfamily: proline-rich protein

Query Match 15.7%; Score 101.5; DB 2; Length 300;
Best Local Similarity 28.8%; Pred. No. 0.91;
Matches 38; Conservative 10; Mismatches 55; Indels 29; Gaps 7;

Qy 4 RGRSTYR-PRPRRYVEPPEMIGP-MRPEQFSDEVFATPEEGEPATQRDPAQAAGEDE 61

Db 121 QGSSOORPPQGNQGGPPPGGPPQPGQGNQGGP--PPPGPPQORPPQPGNGGPPQ 178

Qy 62 GA-----SAGQGPKEAHSQ-----EQGHPTGCECDGPD-----GQEMDPPNP 101

Db 179 GGPHPPPRPGNQGGPPPGGPPQGPQRTQPGNQGGPPQGG-----GPAQPRPGNQGGPPQ 233

Qy 102 EEVKTPEGEKQ 113

```
DB 234 GPOGPPRTGNQQ 245
      | | | | |
RESULT 15
B97586
Hypothetical protein AGR_C_3445 [Imported] - Agrobacterium tumefaciens (strain C58, Cere
C:Species: Agrobacterium tumefaciens
C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 11-Jan-2002
C:Accession: B97586
A.: Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Ourollo, B.; Goldman,
R.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A:Reference number: A97359; PMID:11743194
A:Accession: B97586
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-754 <KUR>
A:Cross-references: GB:AE007869; PIDN:AAK87643.1; PID:g15156995; GSPDB:GN00169
C:Genetics:
A:Gene: AGR_C_3445
A:Map position: circular chromosome

Query Match 15.6%; Score 101; DB 2; Length 754;
Best Local Similarity 27.7%; Pred. No. 2.3;
Matches 36; Conservative 15; Mismatches 47; Indels 32; Gaps 5;

QY 10 RPRPRRVPEPEMIGPMRPEQFSDEVEPATPEEGEPATORDPAAAQGEDEGASACQ-- 67
DB 182 RPRKPKAEQPAGEGEQORPERPRKEPEAK-----EPAAEQQ-PAARPENAEQPAKPREPA 236
QY 68 -GPKPEAHSQEQGHQPQTGCECEDGPDQEMD-----ppNPEEVK 105
DB 237 PGKKPQV--EEKAPEQKAEPAEKPEVPEKKPAAPPEAAKEAPVPTAPTARPPAPEAQ 293
QY 106 TPEEGEKQSQ 115
DB 294 NPAPGQPS 303
```

Search completed: July 1, 2002, 06:33:05
Job time: 227 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 1, 2002, 06:32:39 ; Search time 12.4 Seconds
(without alignments) updates/sec
362.215 Million cell

Title: US-09-782-745-27
Perfect score: 646
Sequence: 1 MSWRGRSTYRPRRYVEPP.....DPPNPEVKTPEGEKQSQ 116

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query Length	ID	Description
1	646	100.0	116	1 GGE2_HUMAN	Q13066 homo sapien
2	625.5	96.8	117	1 GGE7_HUMAN	O76087 homo sapien
3	616.5	95.4	117	1 GGEA_HUMAN	Q13068 homo sapien
4	612.5	94.8	117	1 GGB5_HUMAN	Q13069 homo sapien
5	608.5	94.2	117	1 GGB6_HUMAN	Q13070 homo sapien
6	598	92.6	138	1 GGE1_HUMAN	Q13065 homo sapien
7	589.5	91.3	118	1 GGE3_HUMAN	Q13067 homo sapien
8	266	41.2	146	1 GGB1_HUMAN	O75459 homo sapien
9	120	18.6	102	1 GGC1_HUMAN	O60829 homo sapien
10	100.5	15.6	504	1 TPMS_DROME	P49456 drosophila
11	100.5	15.6	605	1 GLCA_SOYBN	P13916 glycine max
12	99.5	15.4	548	1 NFL_PIG	P02547 sus scrofa
13	99	15.3	839	1 GLT5_WHEAT	P10388 triticum ae
14	98	15.2	838	1 GLT4_WHEAT	P08489 triticum ae
15	97.5	15.1	166	1 PRPC_HUMAN	P02810 homo sapien
16	97.5	15.1	235	1 BAR6_CHITE	O08726 chironomus
17	97.5	15.1	541	1 NFL_RAT	P19527 rattus norv
18	97	15.0	261	1 PRP2_MOUSE	P05142 mus musculu
19	96.5	14.9	229	1 VE4_HPV08	P06425 human papil
20	96.5	14.9	617	1 ABP1_SACEX	P38479 saccharomyc
21	96	14.9	542	1 NFL_MOUSE	P08551 mus musculu
22	96	14.9	554	1 NFL_BOVIN	P02548 bos taurus
23	95.5	14.8	543	1 NFL_HUMAN	P07196 homo sapien
24	95.5	14.8	634	1 SC1_RAT	P24054 rattus norv
25	94.5	14.6	245	1 VE4_HPV05	P06924 human papil
26	94.5	14.6	247	1 PRP4_HUMAN	P10163 homo sapien
27	94	14.6	234	1 PRPM_HUMAN	P10161 homo sapien
28	93.5	14.5	276	1 PRPL_HUMAN	P10162 homo sapien
29	93.5	14.5	580	1 SYN3_HUMAN	O14994 homo sapien
30	93	14.4	174	1 PRPP_HUMAN	P81489 homo sapien
31	93	14.4	206	1 PRP3_RAT	P04474 rattus norv
32	92.5	14.3	183	1 PRPH_MESAU	P06680 mesocricetu
33	92	14.2	466	1 CMGA_RAT	P10354 rattus norv

34	92	14.2	466	1	SX10_MOUSE	Q04888 mus musculu
35	92	14.2	533	1	SX21_MOUSE	O09141 mus musculu
36	92	14.2	595	1	VE1_MCMVS	P11210 murine cyto
37	91.5	14.2	245	1	VE4_HPV5B	P26550 human papil
38	91.5	14.2	579	1	SYN3_RAT	O70441 rattus norv
39	91.5	14.2	601	1	CORO_SCHPO	O13923 schizosacch
40	91.5	14.2	1099	1	IKP4_HSV11	O60721 homo sapien
41	91.5	14.2	1298	1	ICP4_HSV11	P08392 herpes simp
42	91	14.1	417	1	VGLD_HSVBS	Q08100 bovine herp
43	91	14.1	424	1	MR15_DROME	Q9y011 drosophila
44	90.5	14.0	446	1	CMGA_PIG	P04404 sus scrofa
45	90.5	14.0	884	1	ANDR_EULFC	O97776 eulemur ful

ALIGNMENTS

RESULT 1						
GGE2_HUMAN						
ID	GGE2_HUMAN	STANDARD;	PRT;	116 AA.		
AC	Q13066;					
DT	15-JUL-1998 (Rel. 36, Created)					
DT	15-JUL-1998 (Rel. 36, Last sequence update)					
DT	30-MAY-2000 (Rel. 39, Last annotation update)					
DE	GAGE-2 protein (G antigen 2).					
GN	GAGE2.					
OS	Homo sapiens (Human).					
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;					
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
OX	NCBI_TaxID=9606;					
RN	[1]					
RP	SEQUENCE FROM N.A.					
RC	TISSUE=Melanoma;					
RX	MEDLINE=95378788; PubMed=7544395;					
RA	van den Eynde B., Peeters O., de Backer O., Gaugler B., Lucas S.,					
RA	Boon T.;					
RT	"A new family of genes coding for an antigen recognized by autologous					
RT	cytolytic T lymphocytes on a human melanoma."					
RL	J. Exp. Med. 182:689-698(1995).					
CC	-!- FUNCTION: ANTIGEN, RECOGNIZED ON MELANOMA BY AUTOLOGOUS CYTOLYTIC					
CC	T LYMPHOCYTES.					
CC	-!- TISSUE SPECIFICITY: EXPRESSED IN A VARIETY OF TUMOR TISSUES BUT					
CC	NOT IN NORMAL TISSUES, EXCEPT TESTIS.					
CC	-!- SIMILARITY: BELONGS TO THE GAGE FAMILY.					
CC	-----					
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration					
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -					
CC	the European Bioinformatics Institute. There are no restrictions on its					
CC	use by non-profit institutions as long as its content is in no way					
CC	modified and this statement is not removed. Usage by and for commercial					
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/					
CC	or send an email to license@isb-sib.ch).					
CC	-----					
CC	EMBL; U19143; AAA82745.1; -.					
DR	TM; 604244; -.					
KW	Multigene family; Antigen.					
SQ	SEQUENCE 116 AA; 12786 MW; DD305D5CA29AF19A CRC64;					

Query Match 100.0%; Score 646; DB 1; Length 116;
Best Local Similarity 100.0%; Pred. No. 3.8e-38;
Matches 116; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MSWRGRSTYRPRRYVEPP	PMRPFQSFDEVEPATPEEGPATQRPDPAQAQEGED	60
DB	1	MSWRGRSTYRPRRYVEPP	PMRPFQSFDEVEPATPEEGPATQRPDPAQAQEGED	60
QY	61	EGASAGQGPKEAHSQEQGHQPQTGCCECDGPDGQEMDPPNPEVKTPEEGKQSQ	116	
DB	61	EGASAGQGPKEAHSQEQGHQPQTGCCECDGPDGQEMDPPNPEVKTPEEGKQSQ	116	

RESULT 2

```
GGE7_HUMAN
ID GGE7_HUMAN STANDARD; PRT; 117 AA.
AC O76087;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE 30-MAY-2000 (Rel. 39, Last annotation update)
DE GAGE-7 protein (G antigen 7) (GAGE-8) (AL4).
GN GAGE7.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=98316329; PubMed=9651357;
RZ Chen M.E., Lin S.-H., Chung L.W.K., Sikes R.A.;
RT "Isolation and characterization of PAGE-1 and GAGE-7: new genes
RT expressed in the LNCap prostate cancer progression model that share
RT homology with melanoma-associated antigens."
RL J. Biol. Chem. 273:17618-17625(1998).
RN [2]
SEQUENCE FROM N.A.
RA de Backer O.R.Y.;
RT "Structure, chromosomal location and control of expression of the
RT GAGE genes."
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN SOME PROSTATE CANCER TISSUES BUT
CC NOT IN NORMAL PROSTATE TISSUE.
CC -!- SIMILARITY: BELONGS TO THE GAGE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC expressed in the LNCap prostate cancer progression model that share
CC homology with melanoma-associated antigens."
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
DR EMBL; AF058988; AAC25989.1; -
DR EMBL; AF055474; AAC33677.1; -
DR EMBL; AF055475; AAC33678.1; -
DR MIM; 604132; -
KW Multigene family.
SQ SEQUENCE 117 AA; 12978 MW; 365ED8D41361C7AD CRC64;

Query Match 96.8%; Score 625.5; DB 1; Length 117;
Best Local Similarity 97.4%; Pred. No. 9.6e-37;
Matches 114; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 1 MSWRGRSTYR-PRPRRYVEPPMIGPMRPEQFSDEVPATPEGEFATQRPAAQAQGE 59
Db 1 MSWRGRSTYRPRPRRYVQPPMIGPMRPEQFSDEVPATPEGEFATQRPAAQAQGE 60

QY 60 DEGASAGGPKPEAHSQEGHPQTGCECEDGPDGQEMDPPNPEVKTPEGEKQSQ 116
Db 61 DEGASAGGPKPEAHSQEGHPQTGCECEDGPDGQEMDPPNPEVKTPEGEKQSQ 117

RESULT 3
GGE4_HUMAN
ID GGE4_HUMAN STANDARD; PRT; 117 AA.
AC Q13068;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE GAGE-4 protein (G antigen 4).
GN GAGE4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]

Query Match 96.8%; Score 625.5; DB 1; Length 117;
Best Local Similarity 97.4%; Pred. No. 9.6e-37;
Matches 114; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 1 MSWRGRSTYR-PRPRRYVEPPMIGPMRPEQFSDEVPATPEGEFATQRPAAQAQGE 59
Db 1 MSWRGRSTYRPRPRRYVQPPMIGPMRPEQFSDEVPATPEGEFATQRPAAQAQGE 60

QY 60 DEGASAGGPKPEAHSQEGHPQTGCECEDGPDGQEMDPPNPEVKTPEGEKQSQ 116
Db 61 DEGASAGGPKPEAHSQEGHPQTGCECEDGPDGQEMDPPNPEVKTPEGEKQSQ 117

RESULT 3
GGE4_HUMAN
ID GGE4_HUMAN STANDARD; PRT; 117 AA.
AC Q13068;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE GAGE-4 protein (G antigen 4).
GN GAGE4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
```

```
RP SEQUENCE FROM N.A.
RC TISSUE=Melanoma;
RX MEDLINE=95378788; PubMed=7544395;
RA van den Eynde B., Peeters O., de Backer O., Gaugler B., Lucas S.,
RA Boon T.;
RT "A new family of genes coding for an antigen recognized by autologous
RT cytolytic T lymphocytes on a human melanoma."
RL J. Exp. Med. 182:689-698(1995).
CC -!- TISSUE SPECIFICITY: EXPRESSED IN A VARIETY OF TUMOR TISSUES BUT
CC NOT IN NORMAL TISSUES, EXCEPT TESTIS.
CC -!- SIMILARITY: BELONGS TO THE GAGE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
DR EMBL; U19145; AAA82747.1; -
DR MIM; 604246; -
KW Multigene family.
SQ SEQUENCE 117 AA; 12956 MW; 365ED71B2F9DC7AD CRC64;

Query Match 95.4%; Score 616.5; DB 1; Length 117;
Best Local Similarity 96.6%; Pred. No. 3.9e-36;
Matches 113; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

QY 1 MSWRGRSTYR-PRPRRYVEPPMIGPMRPEQFSDEVPATPEGEFATQRPAAQAQGE 59
Db 1 MSWRGRSTYRPRPRRYVQPPMIGPMRPEQFSDEVPATPEGEFATQRPAAQAQGE 60

QY 60 DEGASAGGPKPEAHSQEGHPQTGCECEDGPDGQEMDPPNPEVKTPEGEKQSQ 116
Db 61 DEGASAGGPKPEAHSQEGHPQTGCECEDGPDGQEMDPPNPEVKTPEGEKQSQ 117

RESULT 4
GGE5_HUMAN
ID GGE5_HUMAN STANDARD; PRT; 117 AA.
AC Q13059;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE GAGE-5 protein (G antigen 5).
GN GAGE5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
RC TISSUE=Melanoma;
RX MEDLINE=95378788; PubMed=7544395;
RA van den Eynde B., Peeters O., de Backer O., Gaugler B., Lucas S.,
RA Boon T.;
RT "A new family of genes coding for an antigen recognized by autologous
RT cytolytic T lymphocytes on a human melanoma."
RL J. Exp. Med. 182:689-698(1995).
CC -!- TISSUE SPECIFICITY: EXPRESSED IN A VARIETY OF TUMOR TISSUES BUT
CC NOT IN NORMAL TISSUES, EXCEPT TESTIS.
CC -!- SIMILARITY: BELONGS TO THE GAGE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
```

```
DR EMBL; U19146; AAA82748.1; -.
DR MIM; 604247; -.
KW Multigene family.
SQ SEQUENCE 117 AA; 12924 MW; 234A865E2ECDD06 CRC64;

Query Match          94.8%; Score 612.5; DB 1; Length 117;
Best Local Similarity 95.7%; Pred. No. 7.3e-36;
Matches 112; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

QY 1 MSWRGRSTYR-PPRRYVPEPPMIGPMRPEQFSDEVEPATPEEGEPATQRQDPAAAOEGE 59
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 1 MSWRGRSTYWPRRRYVQPPVEIGPMRPEQFSDEVEPATPEEGEPATQRQDPAAAOEGE 60
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

QY 60 DEGASAGQPKPAHSEQGHQHPQTGCCECDGPDGQEMDPPNPEEVKTPPEEGEKQSQ 116
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 61 DEGASAGQPKPEADSQEQGHQHPQTGCCECDGPDGQEMDPPNPEEVKTPPEEGEKQSQ 117
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

RESULT 5
GGE6_HUMAN
ID GGE6_HUMAN STANDARD; PRT; 117 AA.
AC Q13070;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE GAGE-6 protein (G antigen 6).
GN GAGE6.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Melanoma;
RX MEDLINE=95378788; PubMed=7544395;
RA van den Eynde B., Peeters O., de Backer O., Gaugler B., Lucas S.,
RA Boon T.;
RT "A new family of genes coding for an antigen recognized by autologous
RT cytolytic T lymphocytes on a human melanoma.";
RL J. Exp. Med. 182:689-698(1995).
CC -!- FUNCTION: ANTIGEN, RECOGNIZED ON MELANOMA BY AUTOLOGOUS CYTOLYTIC
CC T LYMPHOCYTES. COMPLETELY SILENT IN NORMAL ADULT TISSUES, EXCEPT
CC TESTIS.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN A VARIETY OF TUMOR TISSUES BUT
CC NOT IN NORMAL TISSUES, EXCEPT TESTIS.
CC -!- SIMILARITY: BELONGS TO THE GAGE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U19147; AAA82749.1; -.
DR MIM; 604248; -.
KW Multigene family.
SQ SEQUENCE 117 AA; 12892 MW; 234A865E3FCCDD06 CRC64;

Query Match          94.2%; Score 608.5; DB 1; Length 117;
Best Local Similarity 94.9%; Pred. No. 1.4e-35;
Matches 111; Conservative 3; Mismatches 2; Indels 1; Gaps 1;

QY 1 MSWRGRSTYR-PPRRYVPEPPMIGPMRPEQFSDEVEPATPEEGEPATQRQDPAAAOEGE 59
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 1 MSWRGRSTYWPRRRYVQPPVEIGPMRPEQFSDEVEPATPEEGEPATQRQDPAAAOEGE 60
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

QY 60 DEGASAGQPKPAHSEQGHQHPQTGCCECDGPDGQEMDPPNPEEVKTPPEEGEKQSQ 116
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 61 DEGASAGQPKPEADSQEQGHQHPQTGCCECDGPDGQEMDPPNPEEVKTPPEEGEKQSQ 117
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

RESULT 6
GGE3_HUMAN
ID GGE3_HUMAN STANDARD; PRT; 118 AA.
AC Q13067;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE GAGE-3 protein (G antigen 3).
GN GAGE3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Melanoma;
RX MEDLINE=95378788; PubMed=7544395;
RA van den Eynde B., Peeters O., de Backer O., Gaugler B., Lucas S.,
RA Boon T.;
RT "A new family of genes coding for an antigen recognized by autologous
RT cytolytic T lymphocytes on a human melanoma.";
RL J. Exp. Med. 182:689-698(1995).
CC -!- TISSUE SPECIFICITY: EXPRESSED IN A VARIETY OF TUMOR TISSUES BUT
CC NOT IN NORMAL TISSUES, EXCEPT TESTIS.
CC -!- SIMILARITY: BELONGS TO THE GAGE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U19147; AAA82749.1; -.
DR MIM; 604248; -.
KW Multigene family.
SQ SEQUENCE 117 AA; 12892 MW; 234A865E3FCCDD06 CRC64;

Query Match          92.6%; Score 598; DB 1; Length 138;
Best Local Similarity 99.1%; Pred No. 8.2e-35;
Matches 108; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSWRGRSTYRPPRRYVPEPPMIGPMRPEQFSDEVEPATPEEGEPATQRQDPAAAOEGE 60
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 1 MSWRGRSTYRPPRRYVPEPPMIGPMRPEQFSDEVEPATPEEGEPATQRQDPAAAOEGE 60
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

QY 61 EGASAGQPKPAHSEQGHQHPQTGCCECDGPDGQEMDPPNPEEVKTPPEE 109
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 61 EGASAGQPKPEADSQEQGHQHPQTGCCECDGPDGQEMDPPNPEEVKTPPEE 109
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

RESULT 7
GGE1_HUMAN
ID GGE1_HUMAN STANDARD; PRT; 138 AA.
AC Q13065;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE GAGE-1 protein (G antigen 1) (M22-F antigen).
GN GAGE1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Melanoma;
RX MEDLINE=95378788; PubMed=7544395;
RA van den Eynde B., Peeters O., de Backer O., Gaugler B., Lucas S.,
RA Boon T.;
RT "A new family of genes coding for an antigen recognized by autologous
RT cytolytic T lymphocytes on a human melanoma.";
RL J. Exp. Med. 182:689-698(1995).
CC -!- FUNCTION: ANTIGEN, RECOGNIZED ON MELANOMA BY AUTOLOGOUS CYTOLYTIC
CC T LYMPHOCYTES. COMPLETELY SILENT IN NORMAL ADULT TISSUES, EXCEPT
CC TESTIS.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN A VARIETY OF TUMOR TISSUES BUT
CC NOT IN NORMAL TISSUES, EXCEPT TESTIS.
CC -!- SIMILARITY: BELONGS TO THE GAGE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U19142; AAA82744.1; -.
DR MIM; 604243; -.
KW Multigene family; Antigen.
SQ SEQUENCE 138 AA; 15418 MW; 37B8F3909EC4B3B2 CRC64;

Query Match          92.6%; Score 598; DB 1; Length 138;
Best Local Similarity 99.1%; Pred No. 8.2e-35;
Matches 108; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSWRGRSTYRPPRRYVPEPPMIGPMRPEQFSDEVEPATPEEGEPATQRQDPAAAOEGE 60
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 1 MSWRGRSTYRPPRRYVPEPPMIGPMRPEQFSDEVEPATPEEGEPATQRQDPAAAOEGE 60
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

QY 61 EGASAGQPKPAHSEQGHQHPQTGCCECDGPDGQEMDPPNPEEVKTPPEE 109
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 61 EGASAGQPKPEADSQEQGHQHPQTGCCECDGPDGQEMDPPNPEEVKTPPEE 109
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
```

RA Boon T.;
RT "A new family of genes coding for an antigen recognized by autologous
RL cytolytic T lymphocytes on a human melanoma.";
RL J. Exp. Med. 182:689-698(1995).
CC -!- TISSUE SPECIFICITY: EXPRESSED IN A VARIETY OF TUMOR TISSUES BUT
CC NOT IN NORMAL TISSUES, EXCEPT TESTIS.
CC -!- SIMILARITY: BELONGS TO THE GAGE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U19144; AA82746.1; -.
DR MIM; 604245; -.
KW Multigene family.
SQ SEQUENCE 118 AA; 12937 MW; D97EBB19E735103 CRC64;

Query Match 91.3%; Score 589.5; DB 1; Length 118;
Best Local Similarity 94.7%; Pred. No. 2.7e-34;
Matches 108; Conservative 3; Mismatches 2; Indels 1; Gaps 1;

QY 4 RGRSTYR-PPRPRRYVEPPMIGMPRPFQSFDEVEPATPEGEPTATQDPAAQAQGEDEG 62
DB 5 RKGSTYYPRPRRYVQPPVIGMPRPFQSFDEVEPATPEGEPTATQDPAAQAQGEDEG 64

QY 63 ASAGGQPKPEAHSEOGHPOTGCECDGPDQEMDPNPEVKTPPEGEKQSQ 116
DB 65 ASAGGQPKPEAHSEOGHPOTGCECDGPDQEMDPNPEVKTPPEGEKQSQ 118

RESULT 8
GGBL_HUMAN
ID GGBL_HUMAN STANDARD; PRT; 146 AA.
AC 075459;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE G antigen family B 1 protein (Prostate-associated gene protein 1)
DE (PAGE-1) (GAGE-9) (AL5).
GN GAGEB1 OR PAGE1 OR GAGE9.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98316329; PubMed=9651357;
RA Chen M.E., Lin S.-H., Chung L.W.K., Sikes R.A.;
RT "Isolation and characterization of PAGE-1 and GAGE-7: new genes
RT expressed in the LNCap prostate cancer progression model that share
RT homology with melanoma-associated antigens.";
RL J. Biol. Chem. 273:17618-17625(1998).
RN [2]
RP SEQUENCE FROM N.A.
RA Strom T.M., Nyakatura G., Hellebrand H., Drescher B., Rosenthal A.,
RA Meindl A.;
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
CC -!- TISSUE SPECIFICITY: PREFERENTIALLY EXPRESSED IN NORMAL MALE AND
CC FEMALE REPRODUCTIVE TISSUES, PROSTATE, TESTIS, FALLOPIAN TUBE,
CC UTERUS, AND PLACENTA, AS WELL AS IN PROSTATE CANCER, TESTICULAR
CC CANCER, AND UTERINE CANCER.
CC -!- SIMILARITY: BELONGS TO THE GAGE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC -!- SIMILARITY: BELONGS TO THE GAGE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF058989; AAC25990.1; -.

DR MIM; 300288; -.
KW Multigene family.
SQ SEQUENCE 146 AA; 16134 MW; 91ABALE3B498DCA2 CRC64;

Query Match 41.2%; Score 266; DB 1; Length 146;
Best Local Similarity 42.0%; Pred. No. 3.2e-12;
Matches 66; Conservative 7; Mismatches 30; Indels 54; Gaps 3;

QY 1 MSWRGSTYRPPRPRRYVEPPMIGMPRPFQSFDEVEPATPEGEPTATQDPAAQAQGEDEG 60
DB 1 MGFLRLIYRRRPMIYVE-----SSESSDE---QPDEVESPTQSDSTPAERED 48
QY 61 EGASAGQ-----GPKPEAHSEQ 78
DB 49 EGASAAQGPPEADSQLVQPGDPTKRVCLRNEMQKLPAGPEPEADSQEQ 108
QY 79 GHPQTGCECDGPDQEMDPNPEVKTPPEGEKQSQ 115
DB 109 VHPKTGCERGDGDVQELGLPNPEVKTPPEDEGQSQ 145

RESULT 9
GGCL_HUMAN
ID GGCL_HUMAN STANDARD; PRT; 102 AA.
AC 060829;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE G antigen family C 1 protein (Prostate-associated gene protein 4)
DE (PAGE-4) (PAGE-1) (JM27) (GAGE-9).
GN GAGEC1 OR PAGE4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98393718; PubMed=9724777;
RA Brinkmann U., Vasmatzis G., Lee B., Yerushalmi N., Essand M.,
RA Pastan I.;
RT "PAGE-1, an X chromosome-linked GAGE-like gene that is expressed in
RT normal and neoplastic prostate, testis, and uterus.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:10757-10762(1998).
RN [2]
RP SEQUENCE FROM N.A.
RA Strom T.M., Nyakatura G., Hellebrand H., Drescher B., Rosenthal A.,
RA Meindl A.;
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
CC -!- TISSUE SPECIFICITY: PREFERENTIALLY EXPRESSED IN NORMAL MALE AND
CC FEMALE REPRODUCTIVE TISSUES, PROSTATE, TESTIS, FALLOPIAN TUBE,
CC UTERUS, AND PLACENTA, AS WELL AS IN PROSTATE CANCER, TESTICULAR
CC CANCER, AND UTERINE CANCER.
CC -!- SIMILARITY: BELONGS TO THE GAGE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF275258; AAF88037.1; -.
DR EMBL; AJ005894; CAA06751.1; -.
DR EMBL; AF238380; AAF62541.1; -.
DR MIM; 300287; -.
SQ SEQUENCE 102 AA; 11153 MW; CE5D07AFBF73301B CRC64;

Query Match 18.6%; Score 120; DB 1; Length 102;
Best Local Similarity 34.5%; Pred. No. 0.02;

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 1, 2002, 06:32:13 ; Search time 21.85 Seconds
(without alignments)
918.417 Million cell updates/sec

Title: US-09-782-745-27
Perfect score: 646
Sequence: 1 MSWGRSTYRPRRYVEPP.....DPPNPEVKTPPEGEKQSQC 116

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues
Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : SPTREMBL19:*
- 1: sp_archaea:*
 - 2: sp_bacteria:*
 - 3: sp_fungi:*
 - 4: sp_human:*
 - 5: sp_invertebrate:*
 - 6: sp_mammal:*
 - 7: sp_mhc:*
 - 8: sp_organelle:*
 - 9: sp_phage:*
 - 10: sp_plant:*
 - 11: sp_rodent:*
 - 12: sp_virus:*
 - 13: sp_vertebrate:*
 - 14: sp_unclassified:*
 - 15: sp_rvirus:*
 - 16: sp_bacteriap:*
 - 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	637	98.6	116	4 Q9UEU5	Q9ueu5 homo sapien
2	266	41.2	146	4 Q9BSST	Q9bss7 homo sapien
3	229.5	35.5	111	4 Q96GT9	Q96gt9 homo sapien
4	119	18.4	1300	12 Q36421	Q36421 alcelaphine
5	117.5	18.2	128	5 Q18563	Q18563 strongyloid
6	116.5	18.0	956	10 Q9LJ64	Q9lj64 arabidopsis
7	115.5	17.9	144	5 Q18564	Q18564 strongyloid
8	112	17.3	504	11 Q91X93	Q91x93 mus musculus
9	111	17.2	1054	11 Q9JMS9	Q9jms9 mus musculus
10	110	17.0	139	11 Q62882	Q62882 rattus norv
11	110	17.0	221	11 Q62881	Q62881 rattus norv
12	107.5	16.6	110	4 Q96GU1	Q96gu1 homo sapien
13	106	16.4	301	11 Q62105	Q62105 mus musculus
14	105.5	16.3	224	12 Q73450	Q73450 human papil
15	105.5	16.3	260	11 Q64306	Q64306 mus musculus
16	105.5	16.3	910	11 Q9JLE8	Q9jle8 mus musculus

17	105	16.3	737	10 Q94IL4	Q94il4 secale cere
18	104.5	16.2	912	11 Q9JLE7	Q9jle7 mus musculus
19	104.5	16.2	1343	12 Q06635	Q06635 bovine herp
20	104	16.1	396	5 Q04151	Q04151 toxoplasma
21	104	16.1	926	3 Q13305	Q13305 pneumocysti
22	102.5	15.9	815	10 Q41553	Q41553 triticum ae
23	102.5	15.9	830	10 Q03872	Q03872 triticum ae
24	102.5	15.9	999	11 Q9JKR6	Q9jkr6 mus musculus
25	102	15.8	480	5 Q27033	Q27033 theileria p
26	102	15.8	614	5 Q94674	Q94674 plasmodium
27	102	15.8	652	11 Q64139	Q64139 mus sp. cbp
28	102	15.8	1560	5 Q26644	Q26644 strongyloce
29	102	15.8	2047	4 Q15019	Q15019 homo sapien
30	101.5	15.7	300	11 Q61888	Q61888 mus musculus
31	101.5	15.7	543	5 Q27043	Q27043 theileria p
32	101	15.6	897	3 Q96VJ1	Q96vj1 pneumocysti
33	100.5	15.6	847	10 Q9XIB6	Q9xib6 arabidopsis
34	100.5	15.6	1927	2 Q54875	Q54875 streptococc
35	100	15.5	1881	16 Q9L7Q2	Q9l7q2 streptococc
36	99.5	15.4	309	4 Q04118	Q04118 homo sapien
37	99.5	15.4	543	10 Q22120	Q22120 glycine max
38	99.5	15.4	605	10 Q94LX2	Q94lx2 glycine max
39	99.5	15.4	623	5 Q9W4A2	Q9w4a2 drosophila
40	99.5	15.4	1110	13 Q91255	Q91255 petromyzon
41	99.5	15.4	1130	11 Q91WB8	Q91wb8 mus musculus
42	99	15.3	667	11 Q35745	Q35745 mus musculus
43	99	15.3	1021	4 Q15451	Q15451 homo sapien
44	99	15.3	1251	4 Q15450	Q15450 homo sapien
45	98.5	15.2	215	12 Q73449	Q73449 human papil

ALIGNMENTS

RESULT	1
Q9UEU5	
ID	Q9UEU5 PRELIMINARY; PRT; 116 AA.
AC	Q9UEU5
DT	01-MAY-2000 (TREMBLrel. 13, Created)
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE	GAGE-8.
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX	NCBI_TaxID=9606;
RP	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE-99323388; PubMed-10397259;
RA	De Backer O., Arden K.C., Boretcll M., Vantomme V., De Smet C.,
RA	Czekay S., Viars C.S., De Plaen E., Brasseur F., Chomez P.,
RA	Van den Eynde B., Boon T., van der Bruggen P.;
RT	"Characterization of the GAGE genes that are expressed in various
RT	human cancers and in normal testis."
RL	Cancer Res. 59:3157-3165(1999).
DR	EMBL; AF055473; AAC33676.1;
SQ	SEQUENCE 116 AA; 12764 MW; DD3052939E66F19A CRC64;

Query Match 98.6%; Score 637; DB 4; Length 116;
Best Local Similarity 99.1%; Pred. No. 5.1e-50;
Matches 115; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	1 MSWGRSTYRPRRYVEPPPEQFSDEVEPATPEGEPTATQROPAAQEGD 60
Db	1 MSWGRSTYRPRRYVEPPPEQFSDEVEPATPEGEPTATQROPAAQEGD 60
QY	61 EGASAGQPKPEAHSEQOQHPTGCECEDGPDGQMDPPNPEVKTPPEGEKQSQC 116
Db	61 EGASAGQPKPEADSEQOQHPTGCECEDGPDGQMDPPNPEVKTPPEGEKQSQC 116

RESULT 2

```
Q9BS7
ID Q9BS7 PRELIMINARY; PRT; 146 AA.
AC Q9BS7;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE G ANTIGEN, FAMILY B, 1 (PROSTATE ASSOCIATED).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=MELANOMA.;
RA Strausberg R.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC004861; AAH04861.1; -.
SQ SEQUENCE 146 AA; 16150 MW; E6C7BA94D898DCB2 CRC64;

Query Match 41.2%; Score 266; DB 4; Length 146;
Best Local Similarity 42.0%; Pred. No. 1.1e-16;
Matches 66; Conservative 7; Mismatches 30; Indels 54; Gaps 3;

QY 1 MSWRGRSTYRPRRYVEPEMIGMPRPEQFSDEVEPATPEGEPTATQRQDPAAQEGED 60
   : : ||| ||| ||| : ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1 MGFLRLIYRRPMIYVE-----SSESSDE---QPDEVESPTQSDSTPAEERED 48
   : : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 61 EGASAGQ-----GPRPEAHSQEQ 78
   |||||
Db 49 EGASAAQGEPEADSQELVQPKTCGELGDPDKRVKVLNRNEQMKLPAEGPEPEADSQEQ 108
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 79 GHPQTGCEGDGQEMDPPNPPEVKTPEGEKQSQ 115
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 109 VHPKTCGERGDPVQELGLPNPEVKTPEDEGQSQ 145
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 3
Q96GT9 PRELIMINARY; PRT; 111 AA.
ID Q96GT9;
AC Q96GT9;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE SIMILAR TO G ANTIGEN 8.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA, AND CHORIOCARCINOMA;
RA Strausberg R.;
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC009232; AAH09232.1; -.
SQ SEQUENCE 111 AA; 12354 MW; 38C93332C5BA0E14 CRC64;

Query Match 35.5%; Score 229.5; DB 4; Length 111;
Best Local Similarity 45.8%; Pred. No. 1.5e-13;
Matches 54; Conservative 15; Mismatches 38; Indels 11; Gaps 3;

QY 1 MSWRGRSTYRPRRYVEPEMIGMPRPEQFSDEVEPA--TPEGEPTATQRQDPAAQEG 58
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 MSWRGRSTYRPRRSLQPPPLICAM-----LEPTDEEPKEKPKPTKSRNPTPDQR 52
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 59 E-DEGASAGQKPEAHSQEQHPOTGCECEDGPDGQEMDPPNPPEVKTPEGEKQSQ 115
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 53 EDDQGAETQVPLEADLQELCQTKTGDCGEGGTDVKGKILPKAHFKMPEAGEGKSQ 110
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 4
Q36421 PRELIMINARY; PRT; 1300 AA.
```

```
Q36421 PRELIMINARY; PRT; 1300 AA.
AC Q36421;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE PUTATIVE IMMEDIATE EARLY PROTEIN.
OS Alcelaphine herpesvirus 1 (wildbeest herpesvirus).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae; Rhadinovirus.
OX NCBI_TaxID=35252;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C500;
RX MEDLINE=97404659; PubMed=9261371;
RA Essner A., Pflanz R., Fleckenstein B.;
RT "Primary structure of the alcelaphine herpesvirus 1 genome.";
RL J. Virol. 71:6517-6525(1997).
DR EMBL; AF005370; AAC58118.1; -.
DR InterPro; IPR000087; Collagen.
SQ SEQUENCE 1300 AA; 128183 MW; 40F9EFD244F34577 CRC64;

Query Match 18.4%; Score 119; DB 12; Length 1300;
Best Local Similarity 35.2%; Pred. No. 0.015;
Matches 45; Conservative 7; Mismatches 44; Indels 32; Gaps 7;

QY 17 VEPPEMIGPMRPEQFSDEVEPATPE--EGE-----PATQRQDPAAQOEG-----E 59
   : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 527 LSGPEGEGPGEGPEGE--GPEGPEGEGPEGPEGPEGPEGPEGPEGPEGPEGPEGPEE 585
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 60 DEGASAGQKPEAHSQEQHPQTGCECEDG-----PDGQEMDPPN-----PERVKT 106
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 586 DEGPEGPEGEGPEGPEGPEGPEGPEGPEGPEGPEGPEGPEGPEGPEGPEGPEGPKG 645
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 107 PEEGEKQS 114
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 646 P-EGECQS 652
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 5
Q18563 PRELIMINARY; PRT; 128 AA.
ID Q18563;
AC Q18563;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE IGG AND IGE IMMUNOREACTIVE ANTIGEN RECOGNIZED BY SERA FROM PATIENTS
DE WITH STRONGYLOIDIASIS (FRAGMENT).
OS Strongyloides stercoralis.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
OC Panagrolaimoidea; Strongyloidea; Strongyloidea.
OX NCBI_TaxID=6248;
RN [1]
RP SEQUENCE FROM N.A.
RA Ramachandran S., Thompson R.W., Gam A.A., Neva F.A.;
RT "A set of recombinant clones for immunodiagnosis of
RT strongyloidiasis.";
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U90353; AAB65141.1; -.
DR InterPro; IPR000087; Collagen.
ET NON_TER 1
SQ SEQUENCE 128 AA; 13497 MW; B3BDFA30BFAE725 CRC64;

Query Match 18.2%; Score 117.5; DB 5; Length 128;
Best Local Similarity 34.5%; Pred. No. 0.0018;
Matches 38; Conservative 15; Mismatches 48; Indels 9; Gaps 6;

QY 6 RSTYRPRRYVE--PPMIGMPRPEQFSDEVEPATPEGE--PATORQDPAAQOEGDEGA 63
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4 RAIVKPKPTTTQAPPEPEGPEGPEGPEGPEGPEGPEGPEGPEGPEGPEGPEGPEG 63
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 64 SAGQGP--KPEAHSQEQHPQTGCECEDGPDGQEMDPPNPPEVKTPEEGEK 112
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

[illegible]

```
RP SEQUENCE FROM N.A.
RX MEDLINE=20573856; PubMed=11124536;
RA Ikegawa S., Sano M., Koshizuka Y., Nakamura Y.;
RT "Isolation, characterization and mapping of the mouse and human PRG4
RT (proteoglycan 4) genes.";
RL Cyogenet. Cell Genet. 90:291-297(2000).
DR EMBL; AB034730; BAA92310.1; -.
DR MGI; MGI:1891344; Prg4.
DR InterPro; IPR002985; Hemopexin.
DR InterPro; IPR002965; P_rich_extensn.
DR InterPro; IPR001212; Somatomedin_B.
DR Pfam; PF01033; Somatomedin_B; 2.
DR PRINTS; PR01217; PRICHEXTENS.
DR SMART; SM00120; HX; 2.
DR SMART; SM00201; SO; 2.
DR PROSITE; PS00024; HEMOPEXIN; UNKNOWN_1.
DR PROSITE; PS00524; SOMATOMEDIN_B; 2.
SQ SEQUENCE 1054 AA; 115992 MW; 4FC64BFA42283235 CRC64;

Query Match 17.2%; Score 111; DB 11; Length 1054;
Best Local Similarity 30.3%; Pred. No. 0.061;
Matches 36; Conservative 17; Mismatches 42; Indels 24; Gaps 6;

QY 7 STYRPRRYVEP-----PENIGMRPEQFS-DEVEPATPEEGEPATQRQDPAAQEG 58
DB 362 TTKKEPTTRKPEPTTPKEPTTPKEPTTPKEPTTPKEPTTPKEPTTPKEPTTPKE- 420
QY 59 EDEGASAGGP---KPEAHSQEQHPQTGCCEGDPGQEMDPNPEVK--TPEEGE 111
DB 421 -----PGPTTPKPEPTTTKEPEPTTTKEPE-STTRKPEPTTPKEPTTPKEPE 470

RESULT 10
Q62882 PRELIMINARY; PRT; 139 AA.
AC Q62882;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE HYPOTHETICAL 15.5 KDA PROTEIN.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=SKELETAL MUSCLE;
RA Geertman R., McMahon A., Sabban E.L.;
RT "Cloning and characterization of cDNAs for novel proteins with
RT glutamic acid-proline dipeptide tandem repeats.";
RL Biochim. Biophys. Acta, Gene Struct. Expr. 1306:147-152(1996).
DR EMBL; U40628; AAB05668.1; -.
DR InterPro; IPR001315; CARD.
DR SMART; SM00114; CARD; 1.
DR PROSITE; PS50209; CARD; 1.
DR Hypothetical protein.
KW Hypothetical protein.
SQ SEQUENCE 139 AA; 15459 MW; 60573DF693989709 CRC64;

Query Match 17.0%; Score 110; DB 11; Length 139;
Best Local Similarity 31.8%; Pred. No. 0.0092;
Matches 34; Conservative 15; Mismatches 36; Indels 22; Gaps 5;

QY 15 RYVEPEPMIGMRPEQFSDEVEPATPEGE---PATQRQDPAAQEGDEGASAGGPKP 71
DB 47 RASEEEIGGP----EDSAVQPTPEPELEAATKGDEPDLEQMEPE-----PEP 95
QY 72 EAHSQEQHPQTGCCEGDPGQEMDPNPEVKTP--EEGEKQSQ 116
DB 96 EVEPEPEPEPEPEPEPE-----PEPEPEREPDFQEGDESGC 136

RESULT 11
Q62882 PRELIMINARY; PRT; 139 AA.
AC Q62882;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE HYPOTHETICAL 15.5 KDA PROTEIN.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=SKELETAL MUSCLE;
RA Geertman R., McMahon A., Sabban E.L.;
RT "Cloning and characterization of cDNAs for novel proteins with
RT glutamic acid-proline dipeptide tandem repeats.";
RL Biochim. Biophys. Acta, Gene Struct. Expr. 1306:147-152(1996).
DR EMBL; U40628; AAB05668.1; -.
DR InterPro; IPR001315; CARD.
DR SMART; SM00114; CARD; 1.
DR PROSITE; PS50209; CARD; 1.
DR Hypothetical protein.
KW Hypothetical protein.
SQ SEQUENCE 139 AA; 15459 MW; 60573DF693989709 CRC64;
```

```
Q62881 PRELIMINARY; PRT; 221 AA.
AC Q62881;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE HYPOTHETICAL 24.6 KDA PROTEIN.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=PHOCHROMACYTOMA, AND BRAIN;
RA Geertman R., McMahon A., Sabban E.L.;
RT "Cloning and characterization of cDNAs for novel proteins with
RT glutamic acid-proline dipeptide tandem repeats.";
RL Biochim. Biophys. Acta, Gene Struct. Expr. 1306:147-152(1996).
DR EMBL; U40627; AAB05667.1; -.
DR InterPro; IPR001315; CARD.
DR SMART; SM00114; CARD; 1.
DR PROSITE; PS50209; CARD; 1.
DR Hypothetical protein.
KW Hypothetical protein.
SQ SEQUENCE 221 AA; 24576 MW; A7661C9040B2CD4D CRC64;

Query Match 17.0%; Score 110; DB 11; Length 221;
Best Local Similarity 31.8%; Pred. No. 0.015;
Matches 34; Conservative 15; Mismatches 36; Indels 22; Gaps 5;

QY 15 RYVEPEPMIGMRPEQFSDEVEPATPEGE---PATQRQDPAAQEGDEGASAGGPKP 71
DB 129 RASEEEIGGP----EDSAVQPTPEPELEAATKGDEPDLEQMEPE-----PEP 177
QY 72 EAHSQEQHPQTGCCEGDPGQEMDPNPEVKTP--EEGEKQSQ 116
DB 178 EVEPEPEPEPEPEPEPE-----PEPEPEREPDFQEGDESGC 218

RESULT 12
Q96GU1 PRELIMINARY; PRT; 110 AA.
AC Q96GU1;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE UNKNOWN (PROTEIN FOR MGC:16481).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RC SEQUENCE FROM N.A.
RA TISSUE=MELANOMA;
RA Strausberg R.;
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC009230; AAB09230.1; -.
SQ SEQUENCE 110 AA; 11777 MW; 6D6AF4563135BB6C CRC64;

Query Match 16.6%; Score 107.5; DB 4; Length 110;
Best Local Similarity 32.7%; Pred. No. 0.012;
Matches 36; Conservative 16; Mismatches 51; Indels 7; Gaps 2;

QY 4 RGRSTYRPRRYVEPEMIGMRPEQFSDEVEPATPEEGEPATQRQDPAAQEGDEGA 63
DB 7 RQSSESGNDQESSQP---VGPVIVQPTTEKR----QEEPEPTDNGIAPSGEIKNEGA 59
QY 64 SAGQGPKEAHSQEQHPQTGCCEGDPGQEMDPNPEVKTPPEGEKQ 113
DB 60 PAVQGTDVFAFOQELALLKTIEDAPGDGPDVREGTLPTFDPTKVLAGEGQ 109
```

```
RESULT 13
Q62105 ID Q62105 PRELIMINARY; PRT; 301 AA.
AC Q62105
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE 31-KDA PROLINE-RICH SALIVARY PROTEIN OF CLONE PUMP125.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86033799; PubMed=3840480;
RA Clements S., Mehancho H., Carlson D.M.;
RT "Novel multigene families encoding highly repetitive peptide
RT sequences: Sequence analyses of rat and mouse proline-rich protein
RT cDNAs.";
RL J. Biol. Chem. 260:13471-13477(1985).
DR EMBL; M11897; AAA40001.1; -.
SQ SEQUENCE 301 AA; 31253 MW; 7F7D99A3F64C6B91 CRC64;

Query Match 16.4%; Score 106; DB 11; Length 301;
Best Local Similarity 29.3%; Pred. No. 0.047;
Matches 39; Conservative 10; Mismatches 54; Indels 30; Gaps 7;

QY 4 RGRSTYR-PRRRYVEPPMIGP-MRPEQFSDEVEPATPEEGEPATQRQDPAAQEGE 61
Db 121 QGSQQRRPQPGNQGGPPQGGPQRRPQGNQGGP--PPGPGPQRRPQGGNQGGPPQ 178
QY 62 GASAG-----QGKPKPEAHSQ-----EQGHPQTGCCECDGPDG-----QEMDP 100
Db 179 GGQGGPRPQGGPPQGGPQGGPQRRPQGNQGGPQGGP-----GPGGPRPQGGPQ 233
QY 101 PEEVKTPEEGEKQ 113
Db 234 GGQGGPRPQGGPQ 246

RESULT 14
Q73450 ID Q73450 PRELIMINARY; PRT; 224 AA.
AC Q73450
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE E4 PROTEIN (FRAGMENT).
GN E4.
OS Human papillomavirus type 76.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=69985;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98118461; PubMed=9454709;
RA Bellus H., Saegling B., Bergmann K., Shamanin V., de Villiers E.M.;
RT "The genomes of three of four novel HPV types, defined by differences
RT of their L1 genes, show high conservation of the E7 gene and the
RT URR.";
RL Virology 240:359-365(1998).
DR EMBL; Y15174; CAA75462.1; -.
FT NON_TER
QY SEQUENCE 224 AA; 23747 MW; 475B9582B7A16846 CRC64;

Query Match 16.3%; Score 105.5; DB 12; Length 224;
Best Local Similarity 27.0%; Pred. No. 0.038;
Matches 33; Conservative 14; Mismatches 48; Indels 27; Gaps 5;

QY 12 RPRRYVEPPMI-----GPMRPEQFSDEVEPATPEEGEPATQRQD---PAAAQEGED 60
```

```
Db 17 RPRSYGDPPTAPFPPTTTPGRRRPPAPQPPPTAPPALAPREHQNSHPKPTPGKEGTE 76
QY 61 E-----GASAGQGPKEAHSQEQGHPQTGCCECDGPDGQEMDPPNPEEVKT 106
Db 77 EKTALQPPPGKRSRGDAPGAPGP-GQPPEQGPSESDPKDPADPD-RDPPENPEDPD 134
QY 107 PE 108
Db 135 PE 136

RESULT 15
Q64306 ID Q64306 PRELIMINARY; PRT; 260 AA.
AC Q64306
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE PROLINE-RICH PROTEIN.
GN PRPMP5 OR PRP MP5.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=BALB/C; TISSUE=PAROTID;
RX MEDLINE=92174915; PubMed=1339347;
RA Layfield R., Bannister A., Pierce E.J., McDonald C.J.;
RT "cDNA clones for mouse parotid proline-rich proteins. mRNA regulation
RT by isoprenaline and the nucleotide sequence of proline-rich protein
RT cDNA MP5.";
RL Eur. J. Biochem. 204:591-597(1992).
DR EMBL; X63005; CAA44734.1; -.
DR EMBL; X63004; CAA44733.1; -.
DR MGD; MGI:1927478; Prmp5.
SQ SEQUENCE 260 AA; 26644 MW; 70BD1468106A3002 CRC64;

Query Match 16.3%; Score 105.5; DB 11; Length 260;
Best Local Similarity 29.5%; Pred. No. 0.045;
Matches 31; Conservative 15; Mismatches 50; Indels 9; Gaps 4;

QY 11 RPRRYVEPPMIGMRPEQFSDEVEPATPEEGEPATQRQDPAAQEGEDGASAGQGP 70
Db 144 PRPGNQGGPPQGGPQGGPQGGPQGGPQGGP--PPQGGP--QRRPP--QPGNQGGPPQGGP 196
QY 71 --PEAHSQEQGHPQTGCCECDGPDGQEMDPPNPEEVKTPEEGEKQ 113
Db 197 GGQPGNQGGPPQGGPQGGPQGGPQGGPQGGPQGGPQGGPQGGPQGGPQGGPQ 241
```

Search completed: July 1, 2002, 06:35:33
Job time: 200 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 1, 2002, 06:27:08 ; Search time 23.84 seconds
(without alignments)
540.460 Million cell updates/sec

Title: US-09-782-745-27

Perfect score: 646

Sequence: 1 MSWRGRSTYRPRRYVEPP.....DPPNPEVKTPPEGEKQSQ 116

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 11073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_032802.*

- 1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
- 2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
- 3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
- 4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
- 5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
- 6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
- 7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
- 8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
- 9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
- 10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
- 11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
- 12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
- 13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
- 14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
- 15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
- 16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
- 17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
- 18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
- 19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
- 20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
- 21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
- 22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	646	100.0	116	AA83159	GAGE2 polypeptide.
2	638	98.8	116	AA83159	GAGE-2 tumour reje
3	623.5	96.5	117	AA83162	GAGE4 polypeptide.
4	619.5	95.9	117	AA83163	GAGE5 polypeptide.
5	616.5	95.4	117	AA83161	GAGE-4 tumour reje
6	615.5	95.3	117	AA83164	GAGE6 polypeptide.
7	613	94.9	118	AA83160	GAGE1 polypeptide.
8	612.5	94.8	117	AA83160	GAGE-5 tumour reje
9	608.5	94.2	117	AA83160	GAGE-6 tumour reje
10	598	92.6	138	AA83160	GAGE-1 tumour reje
11	596.5	92.3	118	AA83161	GAGE3 polypeptide.

12	589.5	91.3	118	19	AA83159	GAGE-3 tumour reje
13	400	61.9	76	21	AA83159	Human secreted pro
14	271.5	42.0	111	22	AA83159	Human polypeptide
15	268.5	41.6	112	22	AA83159	Novel human diagno
16	266	41.2	146	21	AA83159	A human cancer-ass
17	249.5	38.6	106	22	AA83159	Human polypeptide
18	229.5	35.5	111	22	AA83159	Human protein Seq
19	229.5	35.5	111	22	AA83159	Human cell cycle a
20	224	34.7	117	22	AA83159	Human protein Seq
21	155.5	24.1	60	22	AA83159	Novel human diagno
22	141	21.8	115	22	AA83159	Human NOV4 protein
23	132	20.4	249	22	AA83159	Novel human diagno
24	120	18.6	102	21	AA83159	PAGE-4 polypeptide
25	120	18.6	102	21	AA83159	Human secreted pro
26	113.5	17.6	87	21	AA83159	PAGE1 polypeptide.
27	109.5	17.0	111	22	AA83159	Human NOV2 protein
28	108.5	16.8	79	21	AA83159	PAGE3 polypeptide.
29	104.5	16.2	507	14	AA83159	gIV from BHV-1 str
30	104	16.1	89	20	AA83159	Human 5' EST secre
31	103.5	16.0	281	21	AA83159	Streptococcus equi
32	101	15.6	905	18	AA83159	Human p160 polypep
33	101	15.6	1135	18	AA83159	Glycine max antiim1
34	100.5	15.6	605	19	AA83159	Streptococcus pneu
35	100	15.5	258	19	AA83159	Streptococcus pneu
36	100	15.5	392	12	AA83159	Toxoplasma gondii
37	100	15.5	585	19	AA83159	Streptococcus pneu
38	100	15.5	1881	21	AA83159	Streptococcus pneu
39	99.5	15.4	623	22	AA83159	Drosophila melanog
40	99	15.3	667	19	AA83159	BOPI protein. Mus
41	99	15.3	1468	22	AA83159	Drosophila melanog
42	98.5	15.2	543	22	AA83159	Human polypeptide.
43	98	15.2	428	12	AA83159	Toxoplasma gondii
44	98	15.2	1175	22	AA83159	Novel human diagno
45	97.5	15.1	242	22	AA83159	Novel human diagno

ALIGNMENTS

RESULT 1

AA83159
ID AA83159 standard; Protein; 116 AA.
XX
AC AA83159;
XX
DT 24-JUL-2000 (first entry)
XX
DE GAGE2 polypeptide.
XX
KW PAGE-4; MAGE; reproduction; testis; prostate; fallopian tube;
KW uterus; placenta; cancer; major histocompatibility complex; MHC; CTL;
KW cytotoxic T lymphocyte; immune response; antibody; drug delivery;
KW immunoglobulin.

Homo sapiens.

OS

PN WO200012706-A1.

PD 09-MAR-2000.

PF 31-AUG-1999; 99WO-US200046.

PR 01-SEP-1998; 98US-0098993.

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

PI Pastan I, Brinkmann U, Vasmatazis G, Lee B;

XX WPT: 2000-237869/20.

XX Isolated PAGE-4 protein and peptide used for inducing a cytotoxic T
PT lymphocyte response and for raising antibodies which can be used to
PT detect the presence of PAGE-4 in cell samples or body tissues

XX Disclosure; Figure 1a; 63pp; English.

XX PAGE-4 is a gene preferentially expressed in normal male and female

XX reproductive tissues e.g. prostate, testis, fallopian tube, uterus

CC and placenta, as well as in prostate cancer, testicular cancer and

CC uterine cancer. This expression pattern makes it a target for

CC diagnosis and for vaccine based therapy of such neoplasms.

CC An isolated PAGE-4 peptide which induces a cytotoxic T

CC lymphocyte response when bound to a major histocompatibility complex

CC (MHC) class I molecule or the isolated PAGE-4 protein can be used in

CC immunogenic compositions to raise a cytotoxic T lymphocyte response

CC against cells expressing PAGE-4 including cancer cells of the

CC prostate, uterus and testis. The nucleic acids encoding PAGE-4 or

CC PAGE-4 peptide fragments can also be used in these compositions.

CC Antibodies against PAGE-4 and its peptide fragments can be used in

CC detecting the presence of PAGE-4 and PAGE-4 coding sequences in cell

CC samples or body tissues. The presence of PAGE-4 in tissues which are

CC not related to reproduction can be indicative of the spread of

CC cancerous reproductive tissue. PAGE-4 can also be used to raise

CC antibodies which are then used as the targeting group of

CC immunconjugates comprising toxins used in therapeutic applications.

CC This has applications for drug delivery systems. The PAGE

CC polypeptide shares sequence similarity with the GAGE and MAGE family

CC of proteins.

XX Sequence 116 AA;

SQ

Query Match 100.0%; Score 646; DB 21; Length 116;

Best Local Similarity 100.0%; Pred. No. 8.3e-53;

Matches 116; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSWRGRSTYPRPRRYVEPPMIGMRPEQFSDEVEPATPEGEPTAQDPAQAQGED 60

DB 1 msrwgrstyprryryveppmigrpeqfsdevepatpegepatqrdpaaageged 60

QY 61 EGASAGGPKPEAHSQEQGHPTGCECEDGPDQEMDPNPEVKTPPEGEKQSQC 116

DB 61 egasaggpkpeahsqeqghptgcecedgpdqemdpnpveevktpeegeksqsc 116

RESULT 2

AAW47599

ID AAW47599 standard; Protein; 116 AA.

AC AAW47599;

XX

XX 30-JUL-1998 (first entry)

DE GAGE-2 tumour rejection antigen precursor.

XX

XX GAGE tumour rejection antigen precursor; TRAP; tumour;

KW diagnosis; melanoma; antigen; cytolytic T cell clone proliferation;

KW HLA-typing assay.

XX

OS Homo sapiens.

XX

XX Key Location/Qualifiers

FH Misc-difference 38 /note= "Ala encoded by GAG"

FT

FT Misc-difference 39 /note= "Thr encoded by CCT"

FT

XX

XX WO9749417-A1.

XX

XX 31-DEC-1997.

XX

XX 23-JUN-1997; 97WO-US10850.

XX

XX 24-JUN-1996; 96US-0669161.

XX

XX (LUDW-) LUDWIG INST CANCER RES.

XX Boon-Falieur T, Debacker O, Van Den Eynde B;

PI WPI; 1998-076905/07.

XX N-PSDB; AAV18717.

XX

XX Isolated nucleic acid encoding GAGE tumour rejection antigen

PT precursor - processed by HLA-Cw6 molecules into peptides, useful to

PT diagnose melanomas

XX

XX Example 13; Fig 5; 60pp; English.

XX

XX The present sequence represents a GAGE-2 tumour rejection antigen

CC precursor (TRAP). The protein is expressed in a number of tumours. In

CC contrast the only normal tissue which expresses GAGE TRAP protein

CC is testis. Several GAGE TRAPS have been identified (see AAV18717-21).

CC The major difference between these proteins and GAGE-1 is the absence

CC of a stretch of 143 bases located at position 379 to 521 of the GAGE-1

CC TRAP sequence. The rest of the sequences show mismatches at various

CC position, with the exception of GAGE-3 whose 5' end is totally

CC different from the other GAGE cDNAs for the first 112 bases. This

CC region of GAGE-3 cDNA contains a long repeat and a hairpin structure.

CC The antigens can be used to diagnose melanomas, characterised by

CC expression of a TRAP or presentation of a tumour rejection antigen.

CC Antigens shed into blood or urine can be observed and then used to

CC confirm a diagnosis of melanoma using cytolytic T cell clone

CC proliferation methodologies. Other uses for the processed peptides,

CC include HLA-typing assays for, e.g. skin graft or organ transplants.

XX

SQ Sequence 116 AA;

Query Match 98.8%; Score 638; DB 19; Length 116;

Best Local Similarity 99.1%; Pred. No. 4.6e-52;

Matches 115; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSWRGRSTYPRPRRYVEPPMIGMRPEQFSDEVEPATPEGEPTAQDPAQAQGED 60

DB 1 msrwgrstyprryryveppmigrpeqfsdevepatpegepatqrdpaaageged 60

QY 61 EGASAGGPKPEAHSQEQGHPTGCECEDGPDQEMDPNPEVKTPPEGEKQSQC 116

DB 61 egasaggpkpeahsqeqghptgcecedgpdqemdpnpveevktpeegeksqsc 116

RESULT 3

AAW83162

ID AAW83162 standard; Protein; 117 AA.

AC AAW83162;

XX

XX 24-JUL-2000 (first entry)

DT GAGE4 polypeptide.

DE

XX

XX PAGE-4; MAGE; GAGE; reproduction; testis; prostate; fallopian tube;

KW uterus; placenta; cancer; major histocompatibility complex; MHC; CTL;

KW cytotoxic T lymphocyte; immune response; antibody; drug delivery;

KW immunoconjugate.

XX

OS Homo sapiens.

XX

XX WO200012706-A1.

XX

XX 09-MAR-2000.

XX

XX 31-AUG-1999; 99WO-US20046.

XX

XX 01-SEP-1998; 98US-0098993.

XX

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX

XX Pastan I, Brinkmann U, Vasmatazis G, Lee B;

PI

PN WO9749417-A1.
XX 31-DEC-1997.
XX 23-JUN-1997; 97WO-US10850.
XX 24-JUN-1996; 96US-0669161.
XX (LUDW-) LUDWIG INST CANCER RES.
XX Boon-Falleur T, Debacker O, Van Den Eynde B;
XX WPI; 1998-076905/07.
XX N-PSDB; AAV18719.
XX Isolated nucleic acid encoding GAGE tumour rejection antigen
XX precursor - processed by HLA-Cw6 molecules into peptides, useful to
XX diagnose melanomas
XX Example 13; Fig 5; 60pp; English.
XX The present sequence represents a GAGE-4 tumour rejection antigen
XX precursor (TRAP). The protein is expressed in a number of tumours. In
XX contrast the only normal tissue which expresses GAGE TRAP protein
XX is testis. Several GAGE TRAPS have been identified (see AAV18717-21).
XX The major difference between these proteins and GAGE-1 is the absence
XX of a stretch of 143 bases located at position 379 to 521 of the GAGE-1
XX TRAP sequence. The rest of the sequences show mismatches at various
XX positions, with the exception of GAGE-3 whose 5' end is totally
XX different from the other GAGE cDNAs for the first 112 bases. This
XX region of GAGE-3 cDNA contains a long repeat and a hairpin structure.
XX The antigens can be used to diagnose melanomas, characterised by
XX expression of a TRAP or presentation of a tumour rejection antigen.
XX Antigens shed into blood or urine can be observed and then used to
XX confirm a diagnosis of melanoma using cytolytic T cell clone
XX proliferation methodologies. Other uses for the processed peptides,
XX include HLA-typing assays for, e.g. skin graft or organ transplants.
XX Sequence 117 AA;
SQ

Query Match 95.4%; Score 616.5; DB 19; Length 117;
Best Local Similarity 96.6%; Pred. No. 4.6e-50;
Matches 113; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

QY 1 MSWRGRSTYR-PRPRRYVPEPMIGPMRPFQFSDEVEPATPEEGEPATQRDPAAQAQGE 59
Db 1 mswrgstyywprpryvpemigpmrpfqfsdevepatpeegepatqrqdpaaqaqe 60

QY 60 DEGASAGGPKPEAHSQEGHPQTGCCEGDPDQEMDPNPPEVKTPPEEGEKQSQC 116
Db 61 degasagggpkpeadsqeghpgtgccecdgpdgqemdpnppeevktpeegekqsgc 117

RESULT 6
AAY83164
ID AAY83164 standard; Protein; 117 AA.
XX AAY83164;
AC
XX
XX 24-JUL-2000 (first entry)
DT
XX
DE GAGE6 polypeptide.
XX
XX PAGE-4; MAGE; GAGE; reproduction; testis; prostate; fallopian tube;
XX uterus; placenta; cancer; major histocompatibility complex; MHC; CTL;
XX cytotoxic T lymphocyte; immune response; antibody; drug delivery;
XX immunconjugate.
XX Homo sapiens.
XX OS
XX WO200012706-A1.
XX PN
XX

PD 09-MAR-2000.
XX 31-AUG-1999; 99WO-US20046.
XX 01-SEP-1998; 98US-0098993.
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX Pastan I, Brinkmann U, Vasmatazis G, Lee B;
XX WPI; 2000-237869/20.
XX Isolated PAGE-4 protein and peptide used for inducing a cytotoxic T
XX lymphocyte response and for raising antibodies which can be used to
XX detect the presence of PAGE-4 in cell samples or body tissues
XX Disclosure; Figure 1a; 63pp; English.
XX PAGE-4 is a gene preferentially expressed in normal male and female
XX reproductive tissues e.g. prostate, testis, fallopian tube, uterus
XX and placenta, as well as in prostate cancer, testicular cancer and
XX uterine cancer. This expression pattern makes it a target for
XX diagnosis and for vaccine based therapy of such neoplasms.
XX An isolated PAGE-4 peptide which induces a cytotoxic T
XX lymphocyte response when bound to a major histocompatibility complex
XX (MHC) class I molecule or the isolated PAGE-4 protein can be used in
XX immunogenic compositions to raise a cytotoxic T lymphocyte response
XX against cells expressing PAGE-4 including cancer cells of the
XX prostate, uterus and testis. The nucleic acids encoding PAGE-4 or
XX PAGE-4 peptide fragments can also be used in these compositions.
XX Antibodies against PAGE-4 and its peptide fragments can be used in
XX detecting the presence of PAGE-4 and PAGE-4 coding sequences in cell
XX samples or body tissues. The presence of PAGE-4 in tissues which are
XX not related to reproduction can be indicative of the spread of
XX cancerous reproductive tissue. PAGE-4 can also be used to raise
XX antibodies which are then used as the targeting group of
XX immunoconjugates comprising toxins used in therapeutic applications.
XX This has applications for drug delivery systems. The PAGE
XX polypeptide shares sequence similarity with the GAGE and MAGE family
XX of proteins.
XX Sequence 117 AA;
SQ

Query Match 95.3%; Score 615.5; DB 21; Length 117;
Best Local Similarity 95.7%; Pred. No. 5.7e-50;
Matches 112; Conservative 3; Mismatches 1; Indels 1; Gaps 1;

QY 1 MSWRGRST-YRPRRYVPEPMIGPMRPFQFSDEVEPATPEEGEPATQRDPAAQAQGE 59
Db 1 mswrgstyyrprpryvpemigpmrpfqfsdevepatpeegepatqrqdpaaqaqe 60

QY 60 DEGASAGGPKPEAHSQEGHPQTGCCEGDPDQEMDPNPPEVKTPPEEGEKQSQC 116
Db 61 degasagggpkpeadsqeghpgtgccecdgpdgqevdppnppeevktpeegekqsgc 117

RESULT 7
AAY83160
ID AAY83160 standard; Protein; 118 AA.
XX AAY83160;
AC
XX
XX 24-JUL-2000 (first entry)
DT
XX
DE GAGE1 polypeptide.
XX
XX PAGE-4; MAGE; GAGE; reproduction; testis; prostate; fallopian tube;
XX uterus; placenta; cancer; major histocompatibility complex; MHC; CTL;
XX cytotoxic T lymphocyte; immune response; antibody; drug delivery;
XX immunoconjugate.
XX Homo sapiens.
XX OS

Db 53 eedgaaetqvpdleadlqlsagktgcegnpgddqgkilkpkseqfkmpeggdrqpq 110

RESULT 15

ABG05297

ID ABG05297 standard; Protein; 112 AA.

XX ABG05297;

DT 13-FEB-2002 (first entry)

XX Novel human diagnostic protein #5288.

DE Human; chromosome mapping; gene mapping; gene therapy; forensic;

KW food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.

OS

XX WO200175067-A2.

PN

XX 11-OCT-2001.

PD

XX 30-MAR-2001; 2001WO-US08631.

PF

XX 31-MAR-2000; 2000US-0540217.

PR

XX 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

PA

XX Drmanac RT, Liu C, Tang YT;

PI

XX WPI; 2001-639362/73.

DR

XX N-PSDB; AAS69484.

XX

PT New isolated polynucleotide and encoded polypeptides, useful in

PT diagnostics, forensics, gene mapping, identification of mutations

PT responsible for genetic disorders or other traits and to assess

PT biodiversity -

XX

PS Claim 20; SEQ ID No 35656; 103pp; English.

XX

The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic amino acid sequences of the invention.

Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 112 AA;

SQ

Query Match 41.6%; Score 268.5; DB 22; Length 112;

Best Local Similarity 49.2%; Pred. No. 1e-17;

Matches 58; Conservative 14; Mismatches 35; Indels 11; Gaps 3;

Qy 1 MSWRGRSTYRPRRXYRPPMIGMRPEQFSDEVEPA--TPEEGEPATQRQDPAAAAQ- 57

Db 2 miwrgstyrprrrsvpppelligpm-----lepddeeqqqeepptesrdpapgqr 53

Qy 58 GEDEGASAGQGPKEAHSQEQGHPTGTCECEDGPDGQEMDPNPEVKTPEEGEKQSQ 115

Db 54 kedqgsaktqvpdleadlqlsagktgcegnpgddqgkilkpkseqfkmpeggdrqpq 111

Search completed: July 1, 2002, 06:32:09

Job time: 301 sec

